

H I L G A R D I A

*A Journal of Agricultural Science Published by
the California Agricultural Experiment Station*

VOL. 30

JANUARY, 1961

No. 13

STABILITY OF VARIETY RESPONSE TO EXTENSIVE VARIATIONS OF ENVIRONMENT AND FIELD PLOT DESIGN¹

B. J. HOYLE² and G. A. BAKER³

INTRODUCTION

EXTENSIVE field trials of nine strains of Hannchen barley were made by the University of California's Tulelake Field Station in 1957, 1958 and 1959. The results of these tests were subjected to statistical analysis by three different systems.

The purpose of the tests was to investigate the reliability of new field trial procedures, employing design only to provide a complete range of each variety's attributes for the total test area. The ability of these tests to assess relative yielding abilities over a wide range of local environments was evaluated.

No attempt was made to identify areas of uniform response in advance of planting and no block effect was postulated in the analysis. However, when plantings matured, irregular areas of uniform response were made the basis of one system of analysis.

The new procedures are compared with the classical method for field trials utilizing rigid, formalized designs and interpreting results by the analysis of variance with its "experimental error" component.

The information obtained has far-reaching implications for the planning and analysis of field trials and suggests means of streamlining such trials, greatly reducing expenditures of time, effort and money.

FIELD TRIALS Test Materials

Test materials were standard Hannchen and radiation-induced variants of Hannchen, chosen to assure near isogenic constitution while presenting observable yield and appearance differences. Ten varieties were originally selected and furnished by Dr. Tom Ramage of the U.S.D.A., while he was stationed at Davis, California. One variety was dropped from the 1958 and 1959 trials.

¹ Received for publication April 13, 1960.

² Specialist in Field Station Administration and Superintendent of the Tulelake Field Station, Tulelake.

³ Professor of Mathematics and Statistician in the Experiment Station, Davis.

TABLE 1
CODES AND DESCRIPTION OF FIELD TRIALS, TULELAKE, CALIFORNIA
1957, 1958, AND 1959

Year	Code	No. of Replications	No. of varieties	Row length (feet)	No. of rows per variety	Kind of field	Seeding rate	Design
1957	WL	10	10	1	1	Wet	18 seeds	Latin square
	WH1	3	10	4	1	Wet	Normal*	Random block
	WH2	3	10	4	2	Wet	Normal	Random block
	WB	9	10	4	1	Wet	Normal	Paired block
	XR	5	10	8	2	Normal	Normal	Random block
	RL	10	10	1	1	High fertility	18 seeds	Latin square
	RH1	3	10	4	1	High fertility	Normal	Random block
	RH2	3	10	4	2	High fertility	Normal	Random block
	RB	9	10	4	1	High fertility	Normal	Paired block
	PDL	10	10	1	1	Dry, low fertility	18 seeds	Latin square
	PDH1	3	10	4	1	Dry, low fertility	Normal	Random block
	PDH2	3	10	4	2	Dry, low fertility	Normal	Random block
	PDB	9	10	4	1	Dry, low fertility	Normal	Paired block
	CB					Combination for B design for all three fields		
	CL					Combination for Latin square design for all three fields		
	CH2					Combination for H2 design for all three fields		
	CH1					Combination for H1 design for all three fields		
	T57					Combination for all designs for all three fields		
1958	SP	10	9	4	1	Potato	Normal	Uniform rows
	RP	9	9	8	2	Potato	Normal	Latin square
	H1P	9	9	4	1	Potato	Normal	Latin square
	TP	9	9	1	1	Potato	18 seeds	Latin square
	T2P	9	9	1	1	Potato	18 seeds	Latin square
	H2P	9	9	4	2	Potato	Normal	Latin square
	VB	36	9	4	1	Barley	Normal	Uniform rows
	CB8	38	9	4	1	Barley	Normal	Uniform rows
	SB	10	9	4	1	Barley	Normal	Uniform rows
	8RB	9	9	8	2	Barley	Normal	Latin square
	1BB	9	9	4	1	Barley	Double	Latin square
	2XB	9	9	4	2	Barley	Normal	Latin square
	TB	9	9	1	1	Barley	18 seeds	Latin square
	T2B	9	9	1	1	Barley	18 seeds	Latin square
	2BB	9	9	4	2	Barley	Double	Latin square
	T58					Combination of all 1958 trials		
1957 and 1958	T					Combination of all trials for 1957 and 1958.		
1959	Q	4	9	3	1	Barley	Normal	Latin square
	HD	9	9	3	1	Dry	Normal	Latin square
	R	9	9	1	1	Barley	18 seeds	Latin square
	S	9	9	1	1	Barley	18 seeds	Latin square
	GG†	9	9	4	1	Alkali	Normal	Latin square
	KF†	9	9	4	1	Sandy	Normal	Latin square
	N	9	9	3	2	Barley	Normal	Latin square
	AA	9	9	$\frac{1}{4}$	1	Dry	2 seeds	Latin square
	D	9	9	$\frac{1}{4}$	1	Barley	2 seeds	Latin square
	T9	9	9	$\frac{1}{4}$	1	Barley	4 seeds	Latin square
	BB	9	9	$\frac{1}{4}$	1	Dry	2 seeds	Latin square
	P	9	9	4	1	Barley	Normal	Latin square
	XHR	9	9	10	2	Barley	Normal	Latin square

* Normal seeding rate is 90 pounds per acre.

† Of all the plots listed this one was perhaps the poorest agronomically and in appearance. Extreme differences in height and stand were present.

‡ The row length harvested in this plot, due to damage and weed competition, varied from 18 to 48". The values and ranking order shown at other places in this paper were obtained by using the data as it came from the field and not adjusted in any manner. When adjustments were made to a square inch basis, only slight changes were noted.

Observations were made of yield of grain, total weight of straw and grain, and height. Photographs and frequent visual examinations were made.

A complete agronomic history was kept over the three-year period, including much valuable material not pertinent to this particular study.

Designs

Designs employed during the three years are described in table 1.

Included among the designs employed in 1957 was XR, a complete randomized block replicated five times with conventional rod-row construction, representing the classical model for field trials. This trial occupied 2,000 square feet and was grown in a field very near, and similar in background, to R.

Plantings

1957

A piece of land, field R, had been uniformly cropped to barley for four years in preparation for the time when a uniform piece of ground would be needed for test purposes. This area was kept free of roads, ditches and borders. During these four years, the barley crop showed remarkable uniformity in vigor and growth characteristics. No fertilizer was used in any of the four years nor in the 1957 test.

Within this field an area of 1,560 square feet or 1/28 of an acre was selected for test plot use. This area was divided into four complete plot designs each of which contained nine mutant lines and the standard Hannchen. Similar areas of the same size were selected for field W and field PD. Identical designs, shapes and sizes were used in all three fields.

The fields contiguous to R had been rotated between barley and potatoes for four years. A fluctuating water table in the contiguous fields provided a carryover of nutrients from potatoes and left a residue of fertilizer effects, as indicated by the color and yield responses of the barley in R. The plots were of medium fertility and adequate water. Lodging was moderate and growth was ranker than for PD. One irrigation was made but no fertilizer was added.

The PD field was of low fertility and kept rather dry. Plant growth was sparse and erect; maturity was early, and the grain quality was the highest. No irrigations were made and no fertilizer was applied.

The W field was watered excessively, in all, three times, during the growing season. This excessive water caused severe lodging and second growth and greatly lowered the quality of the grain. Fertility in this field was high.

1958

One field used for this year's experiments had been in barley in 1957 and the other in potatoes. The barley field plots were planted May 13-16 and the potato field plots were planted May 1-3. At Tulalake, barley grown on old potato ground is usually later maturing, higher in protein, lodges readily and has more discolored kernels than barley grown on old barley ground.

In general, the plots were more irregular with respect to uniformity of stand and growth than were the trials for 1957. All the plots ripened early and at the same time, with several rainy spells interrupting the harvest from one trial to the next.

TABLE 2
AVERAGE YIELD IN GRAMS FOR NINE VARIETIES OF HANNCHEN
BARLEY FOR THE DESIGNS OF TABLE 1. NUMBER 9 IS NORMAL
AND THE OTHER 8 ARE MUTANTS

No.	Design	Variety average									No. of replica- tions averaged
		1	2	3	4	5	6	7	8	9	
1.....	RB	236.4	192.3	76.3	161.1	213.1	189.9	233.1	135.4	290.4	9
2.....	RL	76.7	74.7	17.6	52.2	103.7	54.2	95.9	38.8	117.0	10
3.....	RH2	444.3	388.6	121.4	264.0	402.7	356.4	397.8	198.9	501.3	3
4.....	RH1	287.5	251.6	77.3	187.7	220.4	202.6	228.6	220.0	285.3	3
5.....	XR	998.7	733.1	450.4	618.8	704.8	769.4	866.6	658.0	826.6	5
6.....	WB	218.4	207.0	86.5	181.9	274.0	222.5	268.1	178.6	305.7	9
7.....	WL	81.0	99.1	32.3	74.8	122.4	79.2	108.9	64.7	116.9	10
8.....	WH2	366.3	468.5	157.4	326.5	479.4	343.7	460.4	220.3	540.0	3
9.....	WH1	247.8	238.1	90.2	182.4	176.5	163.1	273.7	118.1	304.2	3
10.....	PDB	166.2	124.4	60.5	118.4	187.6	109.0	162.8	101.3	218.7	9
11.....	PDL	70.1	64.8	26.0	43.3	90.4	63.3	84.8	38.9	98.4	10
12.....	PDH2	319.3	249.1	140.7	209.0	378.6	224.5	304.1	214.9	326.0	3
13.....	PDH1	159.5	130.0	89.1	112.6	127.6	111.5	165.2	137.5	209.1	3
14.....	CB	207.0	174.6	74.4	153.8	224.9	173.8	221.3	138.4	271.6	27
15.....	CL	75.9	79.6	25.3	56.8	105.5	65.6	96.5	47.5	110.8	30
16.....	CH2	376.6	368.8	139.8	266.5	420.2	308.2	387.4	211.4	455.7	9
17.....	CH1	231.6	206.6	85.5	160.9	174.8	159.1	222.5	158.5	266.2	9
18.....	T57	18335	15943	7049	12796	18118	14712	18694	11781	21287	(weighted total)
19.....	8RB	564.7	609.7	311.8	420.9	555.6	488.0	625.6	457.0	600.0	9
20.....	1BB	151.8	212.3	73.9	130.1	174.3	192.0	183.7	118.9	228.2	9
21.....	2XB	309.8	374.3	135.5	224.3	369.2	306.2	364.3	195.8	432.9	9
22.....	TB	54.6	81.3	24.5	45.5	91.2	58.4	80.6	35.6	92.7	9
23.....	T2B	54.8	92.9	20.5	42.5	90.1	66.8	87.3	33.9	98.1	9
24.....	2BB	286.5	336.4	175.3	188.4	316.8	297.8	341.3	205.0	391.4	9
25.....	CB8	154.5	214.7	69.2	153.5	212.3	115.7	202.8	174.1	274.7	38
26.....	VB	176.2	195.7	102.2	151.5	168.2	181.0	195.4	139.3	199.3	36
27.....	SB	204.3	242.8	80.3	121.8	203.1	138.9	231.7	85.1	268.1	10
28.....	RP	633.9	679.6	320.7	512.0	523.3	609.1	645.4	497.4	652.0	9
29.....	H1P	213.3	214.2	48.0	138.2	247.0	163.7	268.1	103.1	301.4	9
30.....	TP	62.0	103.4	22.9	53.1	111.4	76.9	106.3	44.8	120.9	9
31.....	T2P	49.3	92.8	14.9	49.6	92.5	63.0	116.1	34.1	115.0	9
32.....	H2P	382.6	346.2	158.8	270.7	434.9	352.9	443.6	291.3	501.1	9
33.....	SP	212.0	268.5	41.2	138.4	183.3	155.2	242.3	85.4	322.7	10
34.....	T58	41247	48605	19285	32445	45043	37927	48842	31488	55325	(weighted total)
35.....	T	59582	64548	26334	45240	63161	52638	67536	43268	76612	(weighted total)
35.....	Q	195.5	258.7	125.6	145.2	259.1	228.4	291.9	123.7	309.0	9
36.....	HD	142.1	142.9	90.2	107.1	185.9	130.6	189.0	112.5	177.9	9
37.....	R	103.5	138.9	52.9	86.6	130.3	102.3	144.8	60.4	152.6	9
38.....	S	95.7	126.4	53.6	81.8	121.0	104.3	132.2	57.2	158.2	9
39.....	GG	65.8	65.7	43.9	51.2	80.1	72.2	96.5	43.1	89.7	9
40.....	KF	121.4	114.7	30.1	49.7	130.6	83.4	138.1	75.8	167.2	9
41.....	N	344.1	440.1	305.5	309.4	430.5	391.1	460.1	298.0	480.9	9
42.....	XHR	881.6	786.7	613.8	560.4	692.7	800.0	799.7	628.3	856.1	9
43.....	AA	3.1	4.7	1.0	2.8	3.7	3.2	6.0	1.7	5.0	9
44.....	D	5.0	5.9	2.8	3.0	8.1	6.1	7.8	2.8	8.5	9
45.....	T9	6.9	9.9	3.6	6.7	11.0	9.7	12.0	3.6	13.7	9
46.....	BB	3.2	6.4	2.2	2.7	5.6	4.1	5.0	2.5	8.4	9
47.....	P	247.6	261.2	148.6	204.2	310.7	273.0	345.2	176.9	352.9	9

1959

Two objectives were established for the year.

The first was to contract the limit of plot size to the smallest practicable minimum, using two to four seeds per plot as in plots AA, D, T and BB. In these designs the 9 by 9 Latin square was compressed to four square feet.

The other objective was to grow the varieties under certain extremes not available at the Field Station. Thus plot KF was grown off-station on very sandy soil under extreme stress of weed competition and bird damage and plot GG was grown in an area of severe salinity. On the saline soil, which was sandy, plant height ranged from six inches to two feet and stand from sparse to heavy. A uniform initial stand was achieved in all cases. The remainder of the plots were grown at the Field Station in three fields, all previously cropped to cereal.

One field ("Dry" in table 1) was not irrigated at any time. These plots were slightly irregular in growth characteristics, most notable being the very irregular emergence over a two-week period.

All the plots indicated as grown on "barley" ground were very good agro-nomically. They were sprinkled to germinate and the stand was uniform and regular.

Yield Data

The average yield in grams for the nine strains of barley for each design are given in table 2. Also given are the numbers of replications averaged in each case.

STATISTICAL ANALYSIS OF FIELD TRIALS

I. Factor Analysis

Factor analysis as developed and applied by Cattell (1952), Holzinger and Harman (1941) and Thurstone (1947) and used by Kendall (1957) for the analysis of similar sets of data was selected as being especially appropriate to this particular problem after a great deal of preliminary study and exploratory analysis of the data. The authors have published a previous report (1960) on factor analysis applied to these trials. Factor analysis was applied to the 1957 and 1958 data only. Further tests, conducted in 1959, confirmed the validity of the earlier findings.

Design "T" applies to 1957 and 1958 only, but is used for quickly assessing the 1959 results. Because of the very high correlation coefficients obtained by factor analysis, the yield trials were continued through 1959 for confirmation. Since a simple comparison of the relative yields of the means is sufficient to show their high agreement with T, actual correlation coefficients were not obtained.

Analyses by the SD technique and the game matrix, described later in this paper, were applied to the data for all three years.

The first step in the factor analysis is the calculation of the correlation coefficients between the mean values of the varieties for each of the designs and combinations of designs. These values are given in table 3.

MATRIX OF CORRELATION COEFFICIENTS FOR 35 DESIGNS FOR

	RB	RL	RH2	RH1	XR	WB	WL	WH2	WH1	PDB	PDL	PDH2	PDH1	CB	CL	CH2	CH1
RB	(.982)	.933	.974	.869	.842	.938	.862	.898	.917	.938	.935	.846	.861	.982	.922	.955	.941
RL	.933	(.994)	.916	.753	.671	.951	.960	.958	.849	.967	.980	.922	.786	.971	.994	.975	.854
RH2	.974	.916	(.975)	.861	.845	.900	.865	.919	.893	.896	.931	.852	.758	.946	.915	.975	.991
RH1	.869	.753	.861	(.932)	.869	.788	.730	.738	.798	.781	.726	.744	.800	.833	.747	.820	.993
XR	.842	.671	.845	.869	(.869)	.710	.596	.621	.786	.695	.704	.707	.716	.770	.664	.757	.854
WB	.938	.951	.900	.788	.710	(.979)	.949	.910	.788	.926	.961	.884	.777	.979	.966	.939	.834
WL	.862	.960	.865	.730	.596	.949	(.981)	.962	.724	.888	.942	.892	.655	.922	.981	.949	.773
WH2	.898	.958	.919	.738	.621	.910	.962	(.983)	.863	.876	.937	.824	.686	.917	.966	.969	.832
WH1	.917	.849	.893	.798	.786	.788	.724	.863	(.958)	.815	.811	.677	.843	.860	.819	.868	.954
PDB	.938	.967	.896	.781	.695	.926	.888	.876	.815	(.972)	.939	.931	.838	.972	.945	.935	.861
PDL	.935	.980	.931	.726	.704	.961	.942	.937	.811	.939	(.986)	.917	.700	.967	.986	.970	.814
PDH2	.846	.922	.852	.744	.707	.884	.892	.824	.677	.931	.917	(.922)	.667	.903	.922	.913	.743
PDH1	.861	.786	.758	.800	.716	.777	.655	.686	.843	.838	.700	.667	(.948)	.842	.742	.739	.917
CB	.982	.971	.946	.833	.770	.979	.922	.917	.860	.972	.967	.903	.842	(.987)	.966	.965	.909
CL	.922	.994	.915	.747	.664	.966	.981	.966	.819	.945	.986	.922	.742	.966	(.994)	.977	.824
CH2	.955	.975	.975	.820	.757	.939	.949	.969	.868	.935	.970	.913	.739	.965	.977	(.983)	.878
CH1	.947	.854	.910	.932	.856	.838	.775	.832	.952	.861	.818	.746	.917	.903	.828	.878	(.905)
T57	.988	.961	.975	.876	.839	.957	.913	.924	.897	.949	.959	.905	.831	.987	.957	.980	.934
8RB	.877	.875	.898	.867	.812	.862	.884	.896	.867	.791	.868	.806	.741	.867	.888	.914	.891
1BB	.848	.814	.882	.753	.632	.852	.850	.912	.795	.714	.840	.641	.637	.831	.844	.875	.796
2XB	.923	.946	.951	.781	.694	.923	.942	.983	.862	.865	.955	.824	.720	.927	.960	.975	.865
TB	.837	.946	.866	.660	.539	.900	.971	.979	.770	.844	.941	.831	.625	.882	.965	.942	.748
T2B	.827	.906	.867	.667	.561	.877	.942	.970	.783	.783	.915	.766	.609	.852	.933	.923	.754
2BB	.888	.894	.913	.726	.687	.876	.868	.922	.836	.805	.930	.756	.735	.880	.908	.919	.824
CB8	.790	.869	.737	.772	.484	.837	.873	.858	.760	.838	.792	.737	.811	.839	.859	.818	.822
VB	.895	.824	.918	.835	.794	.861	.846	.900	.881	.734	.839	.687	.948	.857	.847	.895	.875
SB	.888	.915	.921	.754	.698	.815	.862	.948	.927	.836	.893	.777	.747	.866	.903	.937	.891
RP	.841	.714	.869	.886	.844	.766	.736	.798	.852	.643	.722	.603	.666	.776	.733	.814	.881
H1P	.959	.985	.950	.792	.746	.950	.947	.968	.907	.935	.976	.885	.801	.970	.983	.983	.895
TP	.839	.932	.856	.665	.544	.912	.965	.972	.775	.822	.934	.795	.644	.880	.956	.926	.754
T2P	.828	.913	.815	.624	.542	.889	.932	.950	.825	.793	.904	.733	.682	.860	.929	.886	.769
H2P	.960	.961	.928	.818	.770	.989	.932	.908	.828	.939	.976	.901	.827	.986	.968	.954	.879
SP	.916	.885	.925	.814	.715	.829	.834	.938	.958	.822	.858	.704	.800	.877	.873	.917	.930
T58	.930	.939	.933	.839	.642	.924	.934	.967	.723	.889	.926	.805	.437	.930	.946	.957	.915
T	.957	.955	.955	.860	.765	.943	.939	.965	.908	.899	.945	.843	.812	.957	.959	.974	.928

The average of 153 correlation coefficients for the 1957 designs is 0.870. The similar average for 120 correlation coefficients for 1958 designs is 0.899. The average of 288 correlation coefficients for 1957 designs versus 1958 designs is 0.848. (It should be noted that the design T, which is a weighted average of all 1957 and 1958 tests is *not* to be included as either a 1957 test or as a 1958 test.) These results are not unexpected since the 1957 designs were somewhat more diverse than those for 1958. On the other hand, the agronomic situation was more nearly uniform in 1957 than in 1958. The correlations for designs of one year with those of another might be expected to be somewhat smaller than the correlation coefficients for either year because of possible year-variety and/or year-design interactions. These effects appear to be small but statistically significant.

The first factor to be extracted is the one of most interest for this study. This factor may be designated as the "variety difference factor," F_1 . Table 4 ranks the designs with respect to F_1 and also lists F_2 , which indicates the part that differences in varieties contribute to the differences in observed

ESTING 9 VARIETIES OF HANNCHEN BARLEY—TULELAKE 1957-1958

F57	8RB	1BB	2XB	TB	T2B	2BB	CB8	VB	SB	RP	H1P	TP	T2P	H2P	SP	T58	T
988	.877	.848	.923	.837	.827	.888	.790	.888	.888	.841	.959	.839	.828	.960	.916	.930	.957
961	.875	.814	.946	.946	.906	.894	.869	.824	.915	.714	.985	.932	.913	.961	.885	.939	.955
975	.898	.882	.951	.866	.867	.913	.737	.918	.921	.869	.950	.856	.815	.928	.925	.933	.955
876	.867	.753	.781	.660	.667	.726	.772	.835	.754	.886	.792	.665	.624	.818	.814	.839	.860
839	.812	.632	.694	.539	.561	.687	.484	.794	.698	.844	.746	.544	.542	.770	.715	.642	.765
957	.862	.852	.923	.900	.877	.876	.837	.861	.815	.766	.950	.912	.889	.989	.829	.924	.943
913	.884	.850	.942	.971	.942	.668	.873	.846	.862	.736	.947	.965	.932	.932	.834	.934	.939
924	.896	.912	.983	.979	.970	.922	.858	.900	.948	.798	.968	.972	.950	.908	.938	.967	.965
897	.867	.795	.862	.770	.783	.836	.760	.881	.927	.852	.907	.775	.825	.828	.958	.723	.908
949	.791	.714	.865	.844	.783	.805	.838	.734	.836	.643	.935	.822	.793	.939	.822	.889	.899
959	.868	.840	.955	.941	.915	.930	.792	.839	.893	.722	.976	.934	.904	.976	.858	.926	.945
905	.806	.641	.824	.831	.766	.756	.737	.687	.777	.603	.885	.795	.733	.901	.704	.805	.843
831	.741	.637	.720	.625	.609	.735	.811	.948	.747	.666	.801	.644	.682	.827	.800	.437	.812
867	.867	.831	.927	.882	.852	.880	.839	.857	.866	.776	.970	.880	.860	.986	.877	.930	.957
957	.888	.844	.960	.965	.933	.908	.859	.847	.903	.733	.983	.956	.929	.968	.873	.946	.959
980	.914	.875	.975	.942	.923	.919	.818	.895	.937	.814	.983	.926	.886	.954	.917	.957	.974
934	.897	.796	.856	.745	.751	.826	.827	.878	.881	.898	.754	.769	.879	.930	.912	.928	.978
988)	.921	.847	.943	.881	.865	.900	.818	.900	.906	.841	.979	.877	.857	.978	.908	.952	.974
921	(.958)	.867	.929	.874	.895	.905	.813	.938	.918	.914	.926	.883	.877	.897	.907	.956	.958
847	.867	(.954)	.953	.893	.938	.937	.766	.954	.863	.891	.867	.919	.883	.850	.926	.936	.922
943	.929	.953	(.984)	.972	.980	.977	.832	.936	.955	.847	.971	.974	.939	.937	.949	.984	.983
881	.874	.893	.972	(.994)	.988	.929	.843	.850	.914	.724	.943	.994	.957	.898	.877	.942	.934
865	.895	.938	.980	.988	(.993)	.954	.814	.896	.921	.785	.927	.993	.962	.879	.898	.952	.937
900	.905	.937	.977	.929	.954	(.977)	.773	.908	.935	.819	.933	.941	.916	.912	.924	.951	.948
818	.813	.766	.832	.843	.814	.773	(.890)	.737	.807	.674	.847	.850	.840	.834	.836	.890	.879
900	.938	.954	.936	.850	.896	.908	.737	(.972)	.883	.972	.898	.875	.873	.872	.927	.944	.944
906	.918	.863	.955	.914	.921	.935	.807	.883	(.955)	.814	.949	.902	.898	.859	.972	.951	.948
841	.914	.891	.847	.724	.785	.819	.674	.972	.814	(.972)	.808	.753	.751	.787	.880	.880	.881
979	.926	.867	.971	.943	.927	.933	.847	.898	.949	.808	(.985)	.939	.935	.968	.932	.971	.983
877	.883	.919	.974	.994	.993	.941	.850	.875	.902	.753	.939	(.994)	.974	.906	.882	.952	.941
857	.877	.883	.939	.957	.962	.916	.840	.873	.898	.751	.935	.974	(.974)	.890	.887	.941	.926
978	.897	.850	.937	.898	.879	.912	.834	.872	.859	.787	.968	.906	.890	(.989)	.860	.940	.961
908	.907	.926	.949	.877	.898	.924	.836	.927	.972	.880	.932	.882	.887	.860	(.972)	.965	.958
952	.956	.936	.984	.942	.952	.951	.890	.944	.951	.880	.971	.952	.941	.940	.965	(.998)	.998
974	.958	.922	.983	.934	.937	.948	.879	.944	.948	.881	.983	.941	.926	.961	.958	.998	(.998)

mean yields. It is emphasized that the factor analysis is based entirely on the mean yields for the varieties.

Other factors if used could reflect the differences in designs due to sizes of plots, number of replications, year grown and so on.

The first factor has a loading of 0.999 for design T (weighted average of all plots for 1957 and 1958) which means that this design gives a very adequate assessment of the relative yielding abilities of the varieties. The high loadings on several of the designs scattered between the two first years and different fields indicate that the relative yielding ability as determined by some designs is almost independent of the very different fields on which the tests were made and of the quite different years in which the two experiments were conducted.

These remarks are derived from the yield data only. If we consider lodging, grain quality, bundle weights and other characteristics, the ranking of best to worst varieties in many cases may be entirely different.

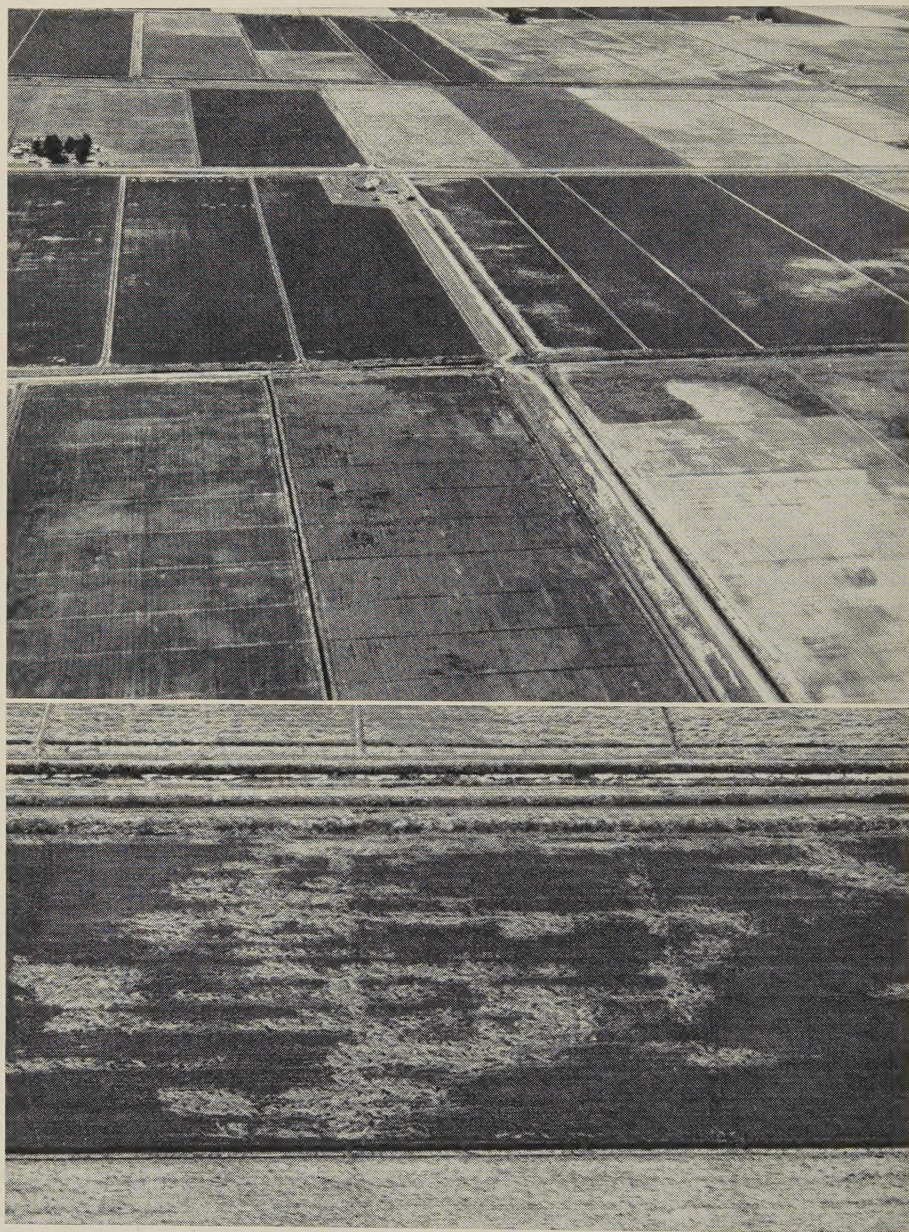


Fig. 1. Air photos of fields at Tulelake showing visible islands of variation. The larger view (top) indicates that irregularities of plant growth are fairly common to all fields and that the more distant fields appear less irregular only because of lack of perception.

The lower photo is a close-up of an individual barley field showing the irregularities of growth habit in more detail. Each sub-area of irregularity (with a common basis) is called an island of variation. None of these islands could have been predicted *before* growth; yet conventional methods hold that random choice of sub-plot location for treatments is best. These prior *assumptions* with further *assumed* row, block, columns, or other effects are held to be inadequate on which to base an analysis, especially when islands of variation are easily calculated. Colwell (1956) has published many similar photos.

STATISTICAL ANALYSIS OF FIELD TRIALS

II. SD Analysis

Examination of nearly all field areas, large or small, will reveal irregular areas of variation. Examples of such sub-areas of irregularity within larger areas are shown in figure 1 where it is further demonstrated that the farther one is from an area, the greater is the apparent lack of irregularity. The areas of irregularity as shown here are examples of one type of "islands of variation." Islands for various attributes have been found present in field areas of less than one square foot in size.

TABLE 4
RANKING OF 35 DESIGNS WITH RESPECT TO THE LOADINGS
FOR THE VARIETY DIFFERENCE FACTOR (F_1).

No.	Design	F_1	F_1^2	No.	Design	F_1	F_1^2
1.....	T	.999	.998	19.....	SB	.946	.895
2.....	H1P	.991	.982	20.....	WL	.943	.889
3.....	CH2	.984	.968	21.....	VB	.940	.884
4.....	T57	.984	.968	22.....	TP	.940	.884
5.....	2XB	.983	.966	23.....	TB	.938	.880
6.....	CL	.972	.945	24.....	T2B	.934	.872
7.....	H2P	.971	.943	25.....	T2P	.923	.852
8.....	T58	.971	.943	26.....	CH1	.922	.850
9.....	CB	.970	.941	27.....	PDB	.920	.846
10.....	RL	.969	.939	28.....	1BB	.908	.824
11.....	WH2	.968	.937	29.....	WH1	.899	.808
12.....	RB	.967	.935	30.....	PDH2	.865	.748
13.....	RH2	.964	.929	31.....	CB8	.864	.746
14.....	PDL	.960	.922	32.....	H1P	.861	.741
15.....	WB	.953	.908	33.....	RH1	.848	.719
16.....	SP	.950	.902	34.....	PDH1	.806	.650
17.....	8RB	.947	.897	35.....	XR	.765	.585
18.....	2BB	.946	.895				

An island of variation is defined as a group of values within which a common level of productivity potential exists regardless of any superimposed treatments. The productivity potential may be measured by yield, protein, height, color, or any other attribute. Areas of similar productivity potential often occur and can be plotted as irregular areas with a connected geographic boundary, because of which they are called islands. Each attribute usually has its own pattern of islands.

Islands of variations can be used to provide a highly reliable analysis of field trials. Such an analysis has been made of the field trials studied for this report and the results confirm those of factor analysis and game analysis.

The elements of this method of analysis based on islands of variation are described in the following basic rules: first, sample points form patterns which are easily determined; second, no sample point is precise; third, each sample point is part of both a range and a continuum; fourth, corresponding values from several treatments can be determined.

Each treatment within a plot has a potential maximum and minimum value extending from areas of the highest to areas of the lowest productivity.

This spread from maximum to minimum is the treatment continuum. The actual sample points form the treatment range. When two or more treatments are studied, they will usually exhibit sample points of similar value occupying the same relative positions on their respective continua. When there is a close relationship between the range and the continuum for each of two or more treatments, then the sample points are said to be in close correspondence. These corresponding values of the various treatments involved are found to form islands when plotted on a field map. When values for yield are so plotted each island is seen to be an area of similar productivity.

Sample Points Are Not Precise

Each sample gathered is a sample point. Each sample point falls at a particular location on the continuum. Theoretically, a large enough number of samples could be gathered so that there would be one sample for every possible position on the continuum. In practice, such a large number is seldom obtained. Each sample is not actually a point, but a composite: several heads of barley or several feet of row, forming a total value. Also, each measurement is necessarily made at the cost of some measurement and damage error. Certain biological changes are very rapid and since all samples cannot be taken at precisely the same time, an individual value will vary slightly from what it would have been had it been taken at a slightly different time. Each sample is, therefore, subject to plus or minus adjustments, no matter how small, leaving it representing a value which should have been just a little higher or lower than recorded. Thus, it is seen that for no common situation encountered in field plot testing can any individual sample be justified as anything but an indication of a segment of the plot continuum.

Since each sample point represents a portion of the continuum, the fewer points there are the larger an area each must represent, if the whole plot is to be included in the sampling. A measure of how well a plot is sampled for each treatment, becomes, therefore, a matter of how well the complete treatment continuum is represented. Each treatment must have very good representation to justify comparisons for any kind of analysis. Thus, it is the relationship between the ranges and the continua which is most important to understand.

Demonstrating the Relationship between Treatment Range and Plot Continuum

The data for figures 2A, B, C, and D are from two treatments of a larger trial. The same four methods were used to obtain the sample points from each treatment.

Method one provided range one of ten sample points, each consisting of a single row one foot long. Method two provided nine sample points which were from a different area but of the same size. These form range two. Range three and four were both of three sample points. The area of each point of range three was four times that of range one. The area of each point of range four was eight times that of range one. For each treatment, then, 25 sample points were obtained and these are used as the best approximation of the continuum for each treatment within this experimental plot.

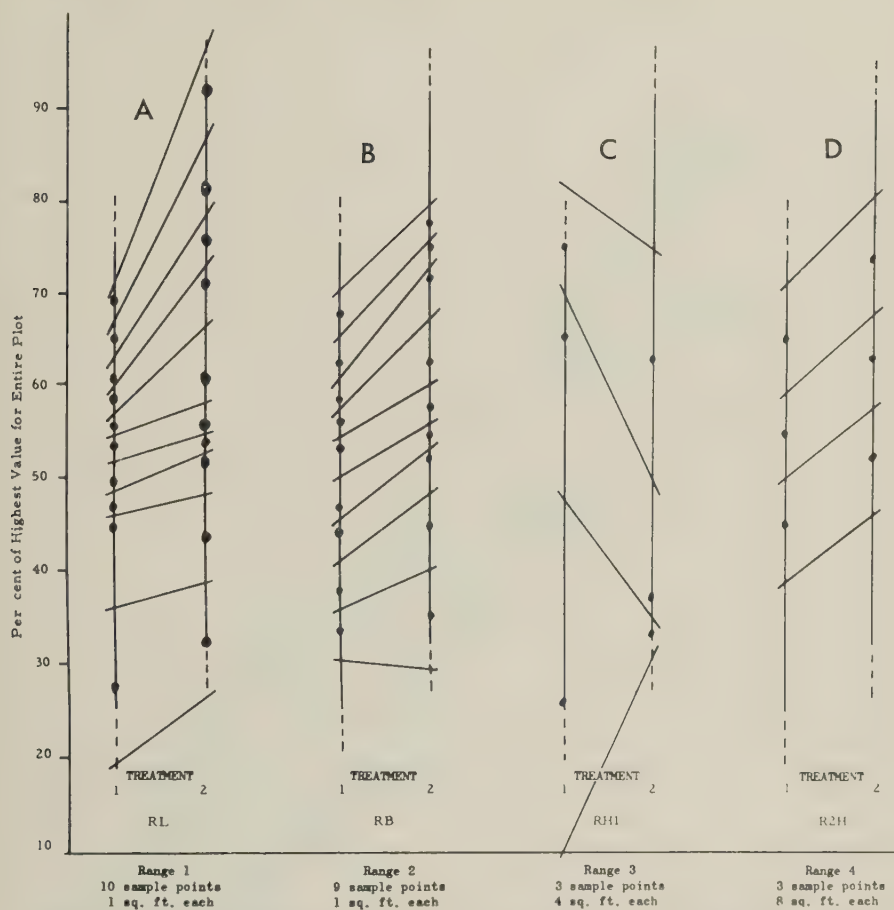


Fig. 2. Showing the relationship between the ranges of two treatments within plots RL, RB, RHI, and R2H and their respective continua. In this case the continua are considered as the known limits of the combined sub-plots covering an area of 1560 square feet.

The solid lines are the known yield continuum and the dotted extensions indicate uncertainty beyond these limits. Each sample point, indicated by a dot, represents a portion of the continuum defined here as that area bounded by the horizontal lines. A plot is adequate only when the corresponding sample points for each treatment adequately represent its respective continuum.

Figure 2A shows the ranges obtained by method one for each treatment, plotted on their respective continua. The other figures show the comparisons for the other ranges. We can see now at a glance how well each range does in representing its respective plot continuum.

In figures 2A, 2B, and 2D, the following conditions are present: both the maximum and the minimum of the range of treatment 2 are higher than that of treatment 1; the mean value of treatment 2 is higher than that of treatment 1; when the position of sample points is compared, it is seen that the highest sample point, second highest and all other points on the continuum

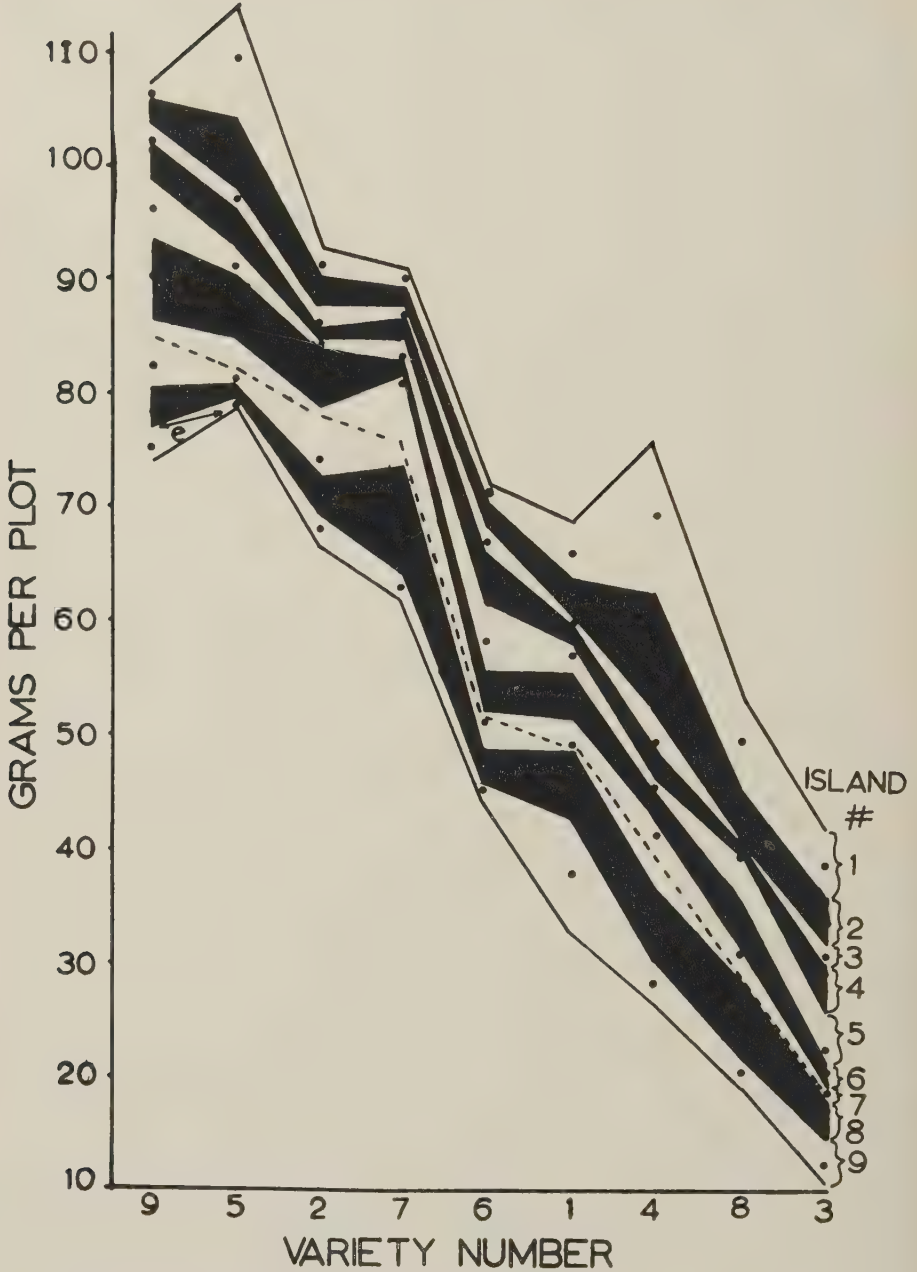


Fig. 3. Showing the correspondence between the sample points of each treatment range for Plot TB when arranged in islands. With perfect correspondence the islands, shown as the diagonal black or white areas, would be parallel. Each sample point more nearly represents the portion of the range indicated by its island width than a precise value. One indication of a treatment being better than another is found when the sample points of the better ones are consistently higher than the corresponding poorer ones. As indicated here, not only the point must be higher, but the entire area it represents must be.

of treatment 2 are higher than their corresponding points on treatment 1. This agreement, in the two treatments, demonstrates good representation. In figure 2C good representation is lacking, as seen by the two ranges failing to meet the three conditions above.

Since a sample point is not definite but represents a portion of the continuum, the exact portion must be defined. The limits, as outlined by the diagonal lines in figure 2, are halfway to the next available point of range. It is assumed that the actual point represents these limits.

Stability of Being Different

The difference between treatments 1 and 2, illustrated in figure 2, was shown to be not only a difference in mean values but—and more importantly—a difference, *in the same direction* of each corresponding value of the two ranges.

It is apparent that each sample point represents an actual area of productivity, the higher the value the more favorable the actual plot conditions. To compare the corresponding points of the two ranges, then, is to compare sub-areas of productivity, within a field plot, which are as nearly equal as possible.

Where there is a consistent relationship between the corresponding points of two treatments and the sub-plots they represent, and that relationship is reflected in the means of the two treatments, they are said to be stable.

For the last three years, practically all of the plots at the Tulalake Field Station have been examined for this stability feature. Many hundreds of examples in published data have also been studied. All of these exhibit some aspect of this feature.

The stability feature is demonstrated in figure 3 for plot TB which consisted of a barley variety trial in a 9 by 9 Latin square design. The varieties are arranged in descending order (best to worst) from left to right, and the range of each is plotted above. Diagonal lines are drawn midway between the sets of corresponding sample points forming nine islands. Alternate islands are darkened for contrast. Because the stability factor is present, figure 3 ranks the varieties adequately.

According to the ranking in figure 3, varieties 8, 4 and 1 are each successively better than 3. Variety 6 is about equal to 1 as indicated by the flatter area. Varieties 7, 2 and 5 are each successively better than 6 or 1 while 9 and 5 are about equal.

For the field plot to be adequate for valid comparisons sample points must be in correspondence, as previously noted. If this has been achieved, the island border lines, from variety to variety, will be approximately parallel. Therefore a test of the effectiveness of a plot is to examine it with respect to this parallelism. Island 7, in figure 3, provides one example. The dotted line drawn within the limits of 7 increases in value from variety 3 through variety 9. Similar lines of ascendancy can be drawn in each island, except at the point in island 9 where it decreases as shown at the arrow labeled "e". For each island lines can be constructed in such a manner as to approach the ideal situation of being parallel. For this reason, it is concluded that this is a very good experimental plot.

TABLE 5
THREE ASPECTS OF THE DATA FOR PLOT TB SHOWN IN FIG. 5

1. VARIETY VALUES IN GRAMS ARRANGED IN DESCENDING ORDER.

	Varieties—Yield								
	1	2	3	4	5	6	7	8	9
	66.2	91.4	37.9	69.1	109.3	71.0	90.0	49.0	105.9
	60.1	89.3	33.3	55.0	99.3	70.9	88.3	41.4	104.9
	60.0	86.4	30.0	48.8	96.9	67.0	87.1	41.2	102.0
	59.8	85.1	29.2	46.6	95.2	65.0	86.0	41.2	102.0
	56.6	84.1	22.0	44.9	90.7	58.0	82.5	41.0	95.8
	53.8	83.5	20.0	44.0	89.3	53.2	81.5	32.2	89.6
	49.1	73.6	18.3	41.4	81.3	50.9	81.4	32.0	82.1
	48.1	71.2	17.2	32.0	79.9	45.2	66.0	22.5	77.6
	38.0	67.5	12.2	27.8	79.1	44.5	62.7	20.0	74.7
Average yield, grams	54.6	81.3	24.5	45.5	91.2	58.4	80.6	35.6	92.7

2. YIELD EXPRESSED AS PERCENT OF HIGHEST VARIETY VALUES.

Arbitrary range ¹ of productivity potential		Percent of highest value							
		100	100	100	100	100	100	100	100
1. 90-100%	{	100	98	87	79	90	99	98	84
		90	95	78	70	88	94	97	84
		90	94	76	67	87	91	95	84
		85	93	57	65	83	81	92	84
2. 75-89%	{	81	92	52	63	81	74	91	66
		74	81	48	60	74	71	90	66
		72	78	45	46	73	63	73	46
		57	74	32	40	72	62	70	41
3. 32-74%	{								

3. ANALYSIS OF VARIANCE AS COMPUTED FROM THE DESIGN OF FIGURE 5.

Variance	D.F.	S.S.	M.S.	F
Total	80	51,922.44		
Rep.	8	960.42	120.05	1.22
Treat.	8	44,633.74	5,579.21	56.52**
Error	64	6,328.26	98.87	
L.S.D. 5% = 9.4				
1% = 12.5				

Note 1.—The plots of 90% to 100% productivity correspond to the diagonally marked ones in Fig. 5, those of 75% to 89% productivity to the black ones and those of 32% to 74% to the white.

** Highly significant F-value.

Figure 3 makes it possible to distinguish among varieties, even where such varieties are very close to each other in rank. For example, it is indicated in figure 3 that variety 6 is slightly better than 1. If 6 is better than 1, then this is a much more precise evaluation than that usually possible by conventional analysis methods.

Utilizing the data plotted on figure 3, it is now possible to determine whether or not the so-called "block effect" actually exists. The conventional analysis method, with the analysis of variance, postulates the existence of such a block effect. Table 5 presents results obtained by analysis of variance, and the relationship between blocks and levels of productivity are shown in figure 4.

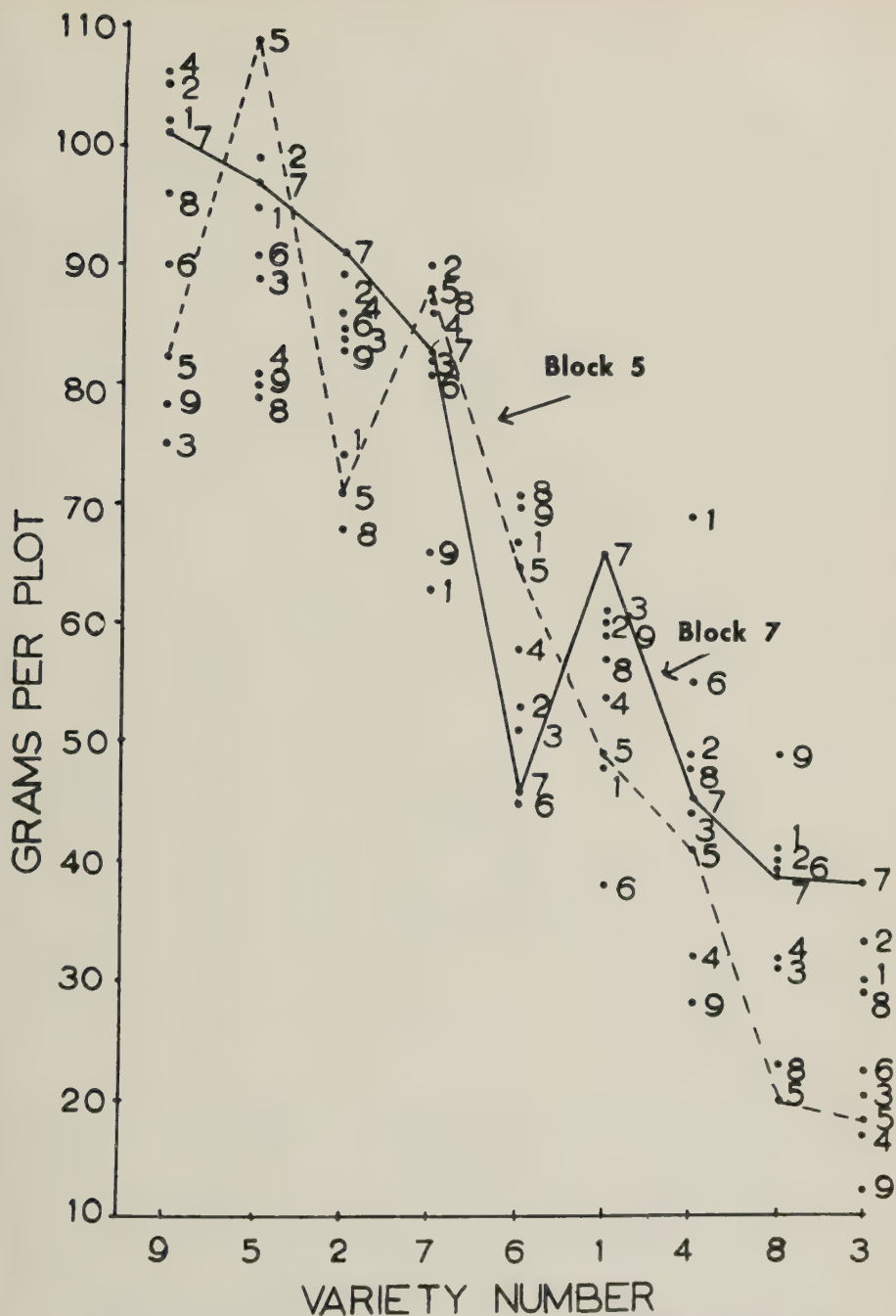


Fig. 4. Showing the correspondence between the sample points of individual treatment ranges and the blocks of Plot TB design from which they were obtained. For a "block effect" to exist, the corresponding sample points from each treatment must be in the same block.

The highest values from each range are said to be in correspondence, the second highest in correspondence, etc. Blocks 7 and 5 are traced as the solid and dotted lines respectively and the complete absence of any block effect is noted. In none of the 41 barley experiments reported here was "block effect" evident.

The data in figure 4 are the same as those of figure 3 with the ranges being plotted in the same manner. In this case, however, each sample point is labeled according to what block it was obtained from. If block effect were present we would expect parallel lines when connecting all the sample points from the same block. However, by connecting the sample points from 2 of the 9 blocks, blocks 7 and 5, we find nothing resembling this parallelism. The solid diagonal line connecting the points from block 7 and the dash-line connecting the points of 5 cross each other at four different points. The same irregularity and criss-crossing would be observable if sample points from all the blocks were connected. It is apparent from this that within-block (or row) comparisons are not valid as such. In this case, varieties 1 and 6 in block 7 are compared at their highest and lowest range values respectively, while in block 5 this trend is drastically reversed.

From the 41 trials reported here and from numerous others, the authors have yet to find a single instance where block effect is peculiar to a given and consistent degree of variation. By contrast, the stability effect as described here is extremely consistent and highly reproducible.

Computing Islands of Variation

Islands of variation have been defined as a clustering together of similar values. Islands of certain types can be observed visually, as illustrated by figure 1, which shows grain lodging and growth types. Islands for protein, yield and other factors must be computed. The boundaries of the sub-plots of similar value are not always connected to each other but may remain isolated. These isolated sub-plots are considered just as much a part of the island of defined productivity as if they were connected. An example of a field plot showing island formation is presented in figure 5 for plot TB which is the same one used in figures 2 and 3 for other purposes.

The method of constructing this map is shown in table 5. Sample points are listed in descending order. Then the percentage of each value of its highest treatment value is determined. Next, arbitrary limits are settled on which in this case are 90 to 100% ; 75 to 89% and 74% and lower. The sub-plots from which each were obtained are plotted on the field map as in figure 5. The island-like groups are clearly evident and also the lack of any block effect.

Other examples of computed islands are shown in figure 6, for the number of heads of lettuce harvested on a certain date, in figure 7 for the yields of single, one-foot rows in a barley plot, in figure 8 for the yield of double eight-foot rows in a barley plot, and in figure 9 plant vigor within a ladino clover trial.

The SD Analysis

One method of using the concepts of islands of variation for an analysis of field plots is the SD analysis. The SD stands for "stability in descending order." It measures the stability of a treatment's rank in each island as compared with the ranking of the mean. Table 6 shows the stability table of the SD technique for the plot TB. From this table conclusions are drawn that treatments 9 and 5 are not different from each other, that 2 and 7 are the

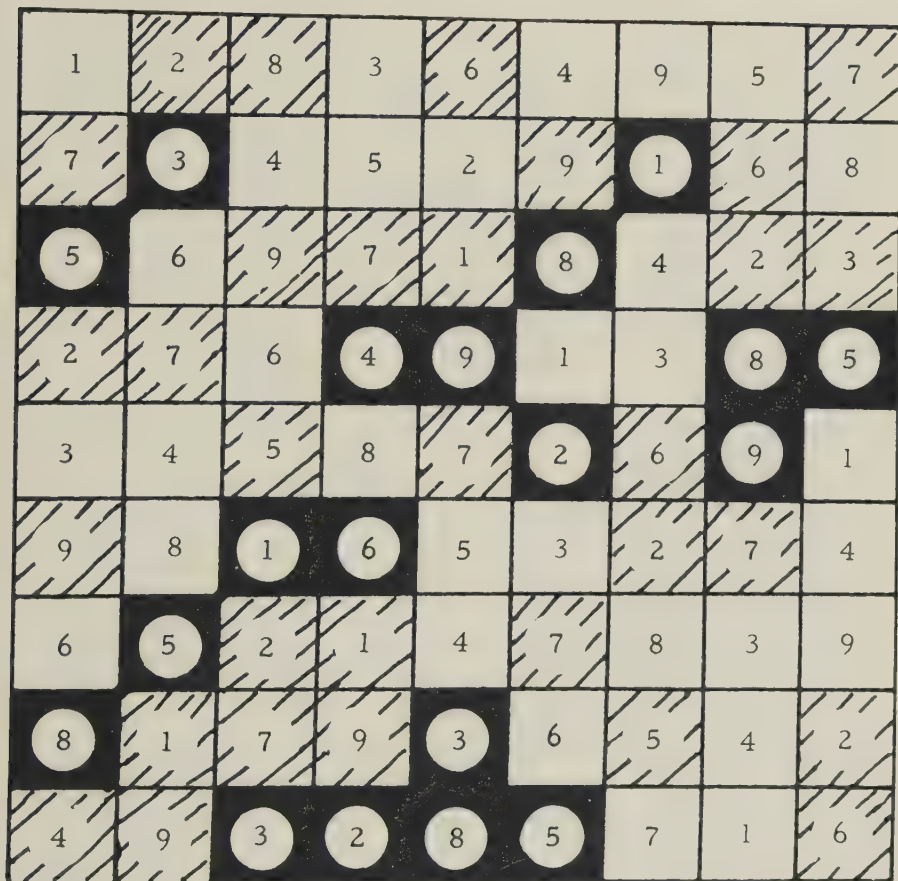


Fig. 5. Calculated islands of variation for yield of Plot TB. The numbers are the variety designation. The sub-plots yielding 90 to 100% of each treatment's highest value are shown under the diagonal lines. Those of 75 to 89% are solid black, and those of 74% and lower are white. Each sub-plot is one row, 1 foot long, of barley.

next lowest in yield and not different, and that treatments 6, 1, 4, 8 and 3 are each successively different from each other; getting progressively worse. This analysis differs from the least significant different groups formed for the 1% level; table 7, where treatments 9, 5, 2 and 7 are not statistically different from each other; 1 and 4 are not different; 4 and 8 are not different and that 8 and 3 are not different. Such an interpretation is confusing and ambiguous. The confusing results arrived at by the analysis of variance for the 28 plots of 1957 and 1958 are shown in table 7.

The results from SD analysis of Plot TB correspond closely with the universal ranking, included as part of table 6. This indicates that the results of the SD analysis are not isolated cases, but provide a valid analysis of past performance on which predictions can be made.

In table 8 the number of experiments ranking the varieties in a given order is shown. This simple grouping supports the conclusions drawn by factor analysis and SD technique.

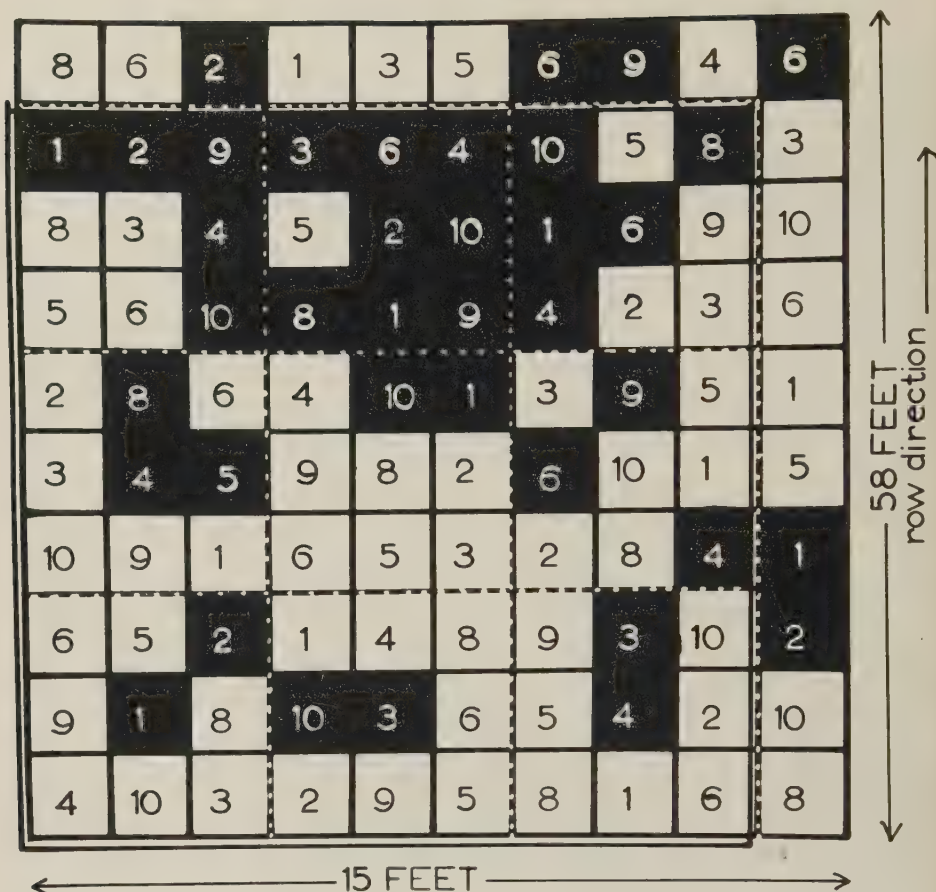


Fig. 6. Calculated islands of variation showing the number of lettuce heads harvested on July 29, 1958. This plot was rated very good agronomically. Each sub-plot contained 5 heads in a 5 foot area. The numbers are variety designations. The white area is where 3 or more heads were harvested and the black where 2 or fewer.

The design is a 9×9 Latin square enclosed within the double solid lines. Furthermore, all 9 varieties are within each sub-group of 9 enclosed by the dotted white lines. To the top, and to the right is an extra row showing that the islands extend indefinitely. Around this entire plot were additional guard rows. This is an interesting design, but is not important except to insure a good distribution.

STATISTICAL ANALYSIS OF FIELD TRIALS

III. Game Theory Applied to Field Trials

The determination of the "proper" variety of cereal to grow may be considered as a "game" with the Grower as one opponent and Nature as the other. An interesting and informative discussion of general game theory, including some reference to games against Nature, has been published by Williams (1954).

A game matrix or rectangular array of the results of various yield trials

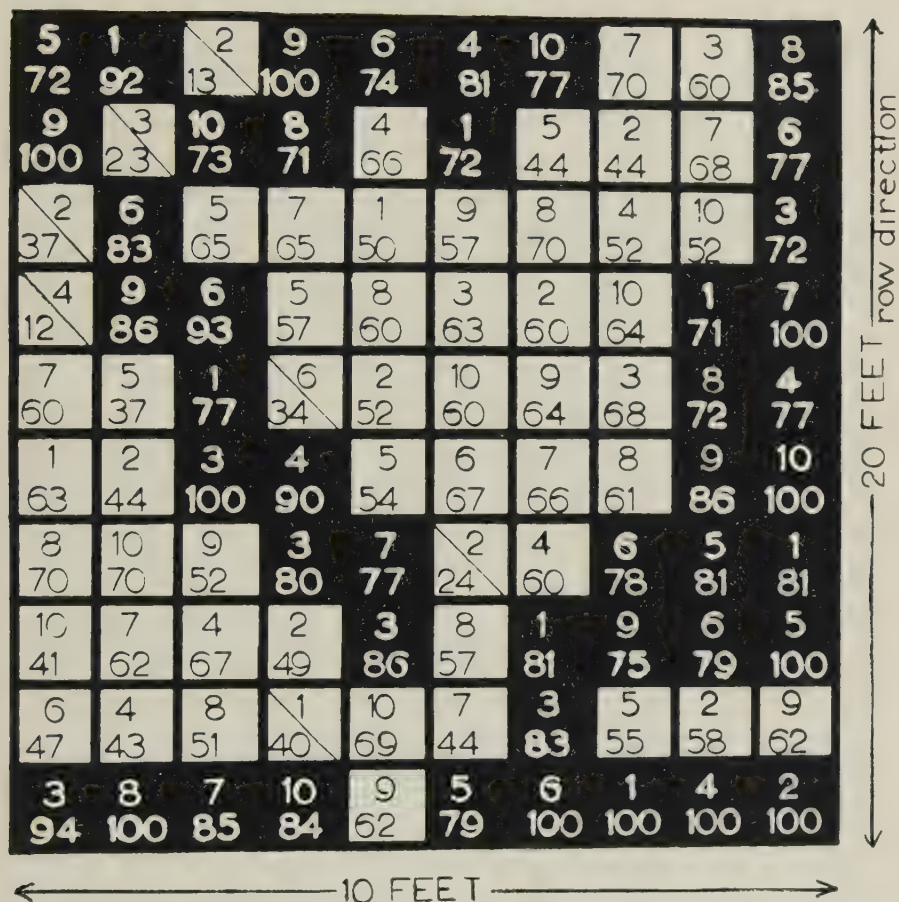


Fig. 7. Showing the calculated islands of variation for yield of Plot PDL. The upper numbers of each sub-plot are variety designation and the lower ones are the per cent of yield for that sub-plot treatment's range. The black area is 71 to 100%, the white area from 41 to 70% and the diagonal area lower than 40%.

for nine strains of Hannechen barley is given in table 9. The entries are the per cent of the general trial mean for each of the nine strains.

A similar game matrix for kernel weights is given in table 10.

Game Matrix for Yields

The totals of the columns in table 9 vary a great deal—from 1904 to 5942. This indicates that some strains yield much higher than others on the average. The reduced sums of squares (the sum of the squares of the deviations from the means of the columns) vary from 36 hundreds to 112 hundreds which shows that some strains are much more erratic in their relative yielding abilities than are others.

The reduced sums of squares for rows show a range of 21 hundreds to 210 hundreds, which means that for some trials there are marked differences

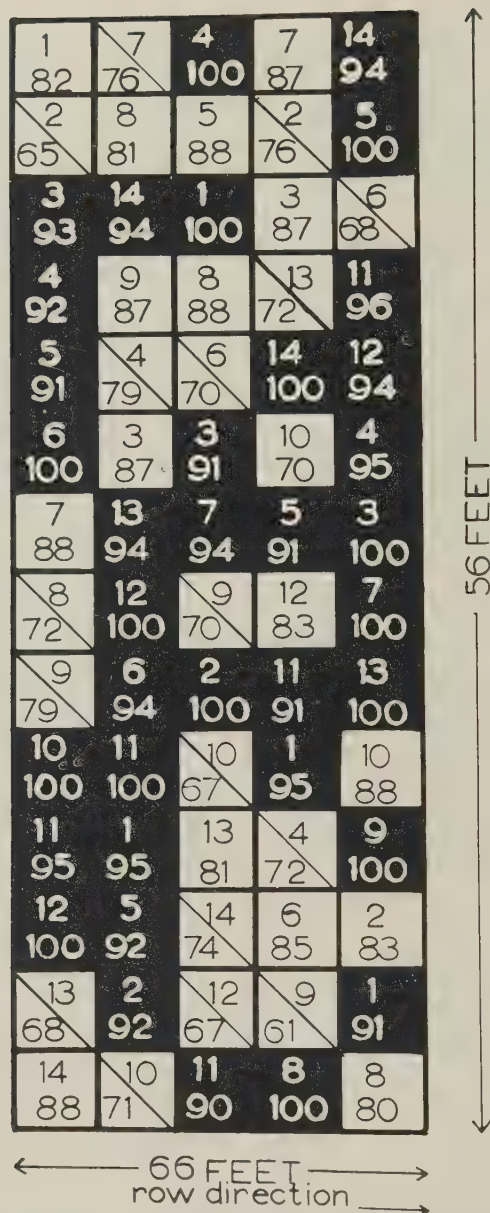


Fig. 8. Calculated islands of variation for yield in a barley plot of 14 feed types in 5 replicated blocks. Each sub-plot is 4 rows 10 feet long. The upper numbers are variety designation and the lower the percentage of that variety's highest yield. Island limits are arbitrary and the one shown here is one possibility. Black squares cover the 90 to 100% range, white squares from 80 to 89% and diagonally marked squares from the lowest values to 79%. Note the lack of block effect.

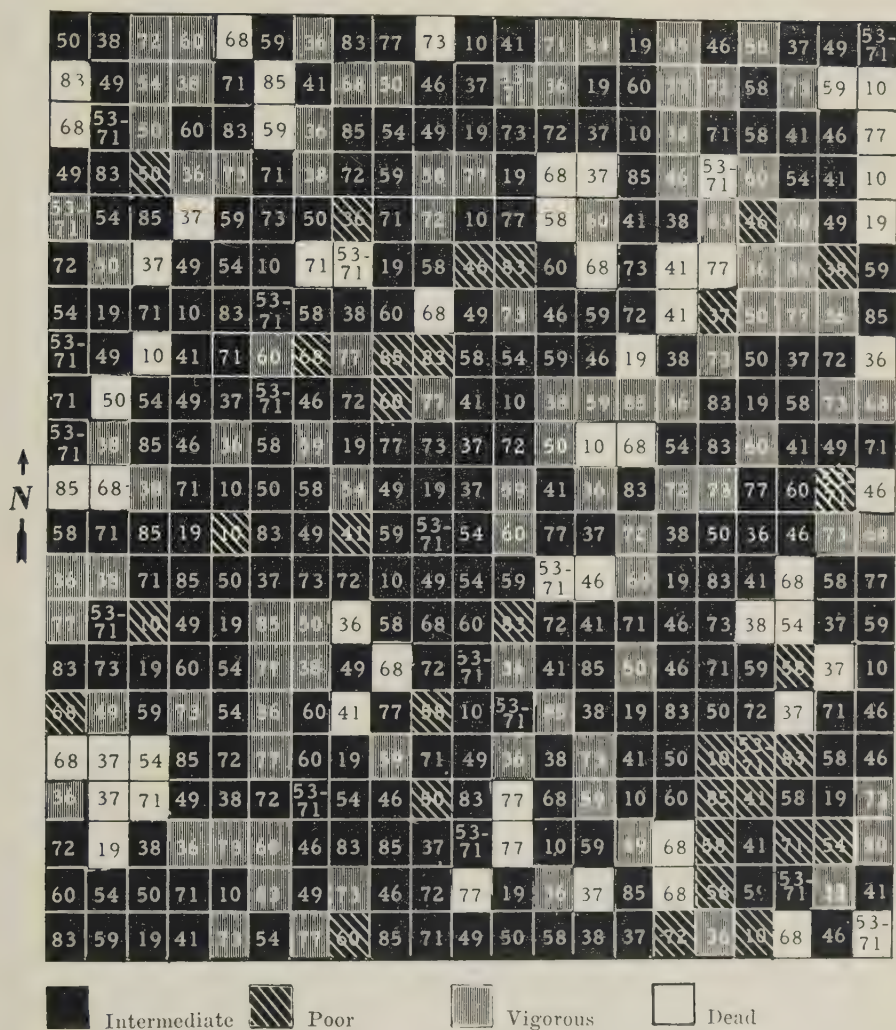


Fig. 9. Island formation in a 21×21 Latin square of ladino clover plot in 1959. Each plant was on a 4 foot spacing and the map shows the survival and vigor after two years of growth. The dead plants are indicated by the white islands. Those plants which are very poor and look as though they may soon die are marked diagonally. The black areas are those of average thrifty growth, while those showing extra vigorous growth and excellent color are marked vertically. Note the absolute lack of row or column effect.

between yielding ability for the strains but in other trials the differences are quite small. Conventional analysis of variance applied to the former trials might find significant differences but completely overlook differences in the latter trials.

We are looking for the best and worst strains or more generally we seek a ranking of the strains and, if possible, an estimate of the superiority of the strains relative to each other.

If we look at the column minima, we find a variation of from 22 to 115. From the standpoint of the Grower we want the minimum for the strain we select to be as large as possible. This is so because then the worst possible result is as good as possible. From this point of view we rank the strains as 9, 7, 2, 5, 1, 6, 4, 8, 3. For comparison, the ranking for T (table 1) is 9, 7, 2, 5, 1, 6, 4, 8, 3, or the agreement happens to be perfect.

TABLE 6

SD ANALYSIS OF THE TB EXPERIMENT OF TABLE 5. THIS SHOWS THE VARIETIES RANKED IN ACCORDANCE WITH THE ISLAND THEORY

Rank	Island number									Mean numerical rank	Mean numerical average	SD* score	Universal rank, T
	1	2	3	4	5	6	7	8	9				
1.....	5	9	9	9	9	9	9	5	5	9	grams 92.7	0	9
2.....	9	5	5	5	5	5	7	9	9	5	91.2	1	7
3.....	2	2	7	7	2	2	5	2	2	2	81.3	0	2
4.....	7	7	2	2	7	7	2	7	7	7	80.6	1	5
5.....	6	6	6	6	6	6	6	1	6	6	58.4	1	1
6.....	4	1	1	1	1	1	1	6	1	1	54.6	2	6
7.....	1	4	4	4	4	4	4	4	4	4	45.5	1	4
8.....	8	8	8	8	8	8	8	8	8	8	35.6	0	8
9.....	3	3	3	3	3	3	3	3	3	3	24.5	0	3

* Measures stability within a group. Zero indicates no deviation. See Appendix A.

SUCCESSIVE GROUPINGS OF TREATMENTS FROM BEST TO WORST OBTAINED BY:

SD analysis : (9,5) (2,7) (6) (1) (4) (8) (3)

Apparent from figure 3 : (9,5) (2) (7) (6,1) (4) (8) (3)

Analysis of variance, table 5 : (9, 5, 2, 7) (1, 4)
(4, 8)
(8, 3)

If we look at the entries in table 9, we see that no strain is always the best and that no strain is always the worst. That is, as we encounter quite different environments in extended time and space we have varying relative responses of these strains as indicated by yield. Sometimes the relative responses are close together, sometimes far apart. Sometimes certain strains appear superior and sometimes others are better. It is the strategy of the Grower to select that strain or combinations of strains that will ensure him the greatest possible return in yield of barley against all the conceivable vagaries (strategies) of Nature. Of course, our knowledge of the vagaries of Nature consists only of the yields that we have obtained in the trials that we have made.

If one strain is always poorer than all the others in a great number of trials, the grower would be inclined never to grow it. This is true of strain 3 except that strain 8 is less for trial 29 (57 vs. 59). Thus, we can say that almost all strains always dominate (exceed) strain 3, and we can then eliminate strain 3 from the game matrix and assign it the rank 9.

If one strain is always better than all the others, we can eliminate it from the matrix and assign it the rank 1. Strain 9 almost fulfills this condition. Actually strain 1 exceeds strain 9 in trials 4, 5, and 36; strain 2 exceeds strain

TABLE 7

NINE VARIETIES OF HANNCHEN BARLEY RANKED ACCORDING TO THEIR MEANS FOR THE FIRST TWENTY-EIGHT TRIALS OF TABLE 1. THE RANKS ONE TO NINE ARE GIVEN ALONG THE TOP AND THE VARIETIES HAVING EACH RANK FOR EACH EXPERIMENT ARE LISTED IN THE BODY OF THE TABLE. GROUPING LINES INDICATE LEAST SIGNIFICANT DIFFERENCE LIMITS FROM ANALYSIS OF VARIANCE AND * INDICATES NONSIGNIFICANCE BY THE SD SCORE

Plot	RANK									Significance by analysis of var.
	1	2	3	4	5	6	7	8	9	
T	9	7	2	5	1	6	4	8	3	
H1P	9	7	5	2	1*	6	4	8	3	1% level
2xB	9	2*	5	7*	1	6	4	8	3	"
H2P	9	7*	5	2	1	6*	8	4	3	"
RL	9	5	7	1	2	6	4	8	3	"
WH2	9*	5	2	7*	1*	6	4*	8	3	"
RB	7	2	9*	1*	5	6	8	4	3	"
RH2	9	1*	5*	8	2	6	4	8	3	"
PDL	9	5*	7*	1	2*	6	4	8	3	"
WB	9*	5	7	6	1	2	4	8	3	"
SP	9	2	7	1	5	6	4	8	3	"
8RB	9	1	7	5	2	6*	4	8*	3	"
2BB	9	7*	2	5*	6	1	8	4	3	"
SB	9	2	7	1	5	6	4	3	8	"
WL	5*	9*	7	2*	1*	6*	4*	8	3	"

Continued Next Page.

Analysis of variance was made according to Snedecor (1956). Each trial was analyzed as a complete randomized block since neither agronomically nor upon investigation by the island method was a row or column gradient found.

Code T which is the weighted mean of all the trials represents the best information available concerning the over-all performance of these varieties. An individual experiment producing a ranking order closest to T can be said to be the best. Use of factor analysis provides another way of making this comparison.

TABLE 7 Continued

Plot	RANK									Significance by analysis of var.
	1	2	3	4	5	6	7	8	9	
VB	(9*)	(2* 7*)	(6*)	(1*)	(5*)	(4*)	8)	(3)		1% level
TP	9	5	7	2	6	1	4	8	3	N.S.
TB	(9	5	2	7)	(6	(1)	(4)	(8)	3)	1% level
T2B	(9	2	5	7)	(6	(1)	(4)	(8)	3)	"
T2P	(7* 9*)	(2	5)	(6	(4	1)	8)	(3)		"
PDB	(9* 5	(1	7)	2	(4	6*)	8	3)		"
1BB	(9	(2	(6	7* 5*)	1)	(4*)	8*)	3)		"
WH1	(9	7* (1	(2*)	4	5* 6*	8)	3)			"
PDH2	(5	(9* (1	7*)	6	(2)	8	4)	3)		"
CB8	(9)	(2* 5	7*)	(8	(1)	4)	(6)	(3)		"
RP	2*	9*	7	1	6	5	4	8*	3	N.S.
RH1	(1	9)	(2* (7*	5* 6)	4)	(8*	3)			5% level
PDH1	9	7*	1*	8	2*	5*	4*	6*	3*	N.S.
XR	1	(7	9	(6	2)	5	8	(4)	3)	1% level

TABLE 8
SHOWING THE NUMBER OF EXPERIMENTS IN WHICH THE GIVEN
VARIETY IS IN A GIVEN RANK—1957-58

Varieties listed according to T	Rank								
	1	2	3	4	5	6	7	8	9
9.....	21	5	2						
7.....	2	7	10	9					
2.....	1	8	5	5	7	2			
5.....	2	7	6	3	5	5			
1.....	2	2	4	6	7	6	1		
6.....	1	3	7	13	2	2	
4.....	1	2	21	5	
8.....	1	1	..	5	20	1
3.....	1	27

From the frequency of the placings, 9 is better than 7 and 2, 5 and 1 are about equal but worse than 7. Also that 6, 4, 8 and 3 are successively worse than the group 2, 5, and 1. Since 2, 5 and 1 are nearly equal we would expect their ranking order to vary from experiment to experiment which is the case as seen in Table 7. For this reason, Factor Analysis is conservative in its results.

TABLE 9

GAME MATRIX FOR GROWER VERSUS NATURE. FORTY-ONE TRIALS ON NINE STRAINS OF HANNCHEN BARLEY (1957-1958-1959). RESULTS IN PERCENT OF TRIAL MEANS

Trial number	Yield Grower (Strains)									Reduced sum of squares (hundreds)
	1	2	3	4	5	6	7	8	9	
1.....	123	100	40	83	110	100	121	71	152	85
2.....	110	107	26	74	148	77	137	56	167	169
3.....	131	114	35	77	117	105	117	58	147	105
4.....	132	116	35	86	101	94	105	101	130	67
5.....	136	98	60	82	98	104	120	87	115	40
6.....	102	96	41	83	126	104	124	83	141	70
7.....	94	114	37	87	141	91	126	75	135	87
8.....	98	126	43	87	128	92	123	60	145	91
9.....	124	120	46	92	88	82	136	60	152	100
10.....	121	89	43	86	135	78	118	72	158	102
11.....	109	101	40	67	132	98	132	61	152	116
12.....	121	94	54	79	142	85	115	88	124	60
13.....	117	93	64	82	93	82	119	99	152	54
14.....	110	119	60	82	110	94	122	85	119	36
15.....	93	130	44	81	108	118	113	74	140	73
16.....	104	125	45	74	122	101	122	64	143	85
17.....	88	129	38	73	145	92	129	57	148	124
18.....	85	143	31	65	138	103	134	52	151	156
19.....	102	119	62	68	113	104	121	73	138	56
20.....	89	123	39	87	121	66	117	101	158	98
21.....	105	117	62	90	100	107	117	83	119	28
22.....	116	139	46	69	116	80	133	48	153	129
23.....	114	119	57	91	94	108	114	88	116	32
24.....	112	114	25	74	131	87	142	55	159	149
25.....	80	132	30	68	143	99	136	58	155	153
26.....	70	133	22	72	132	90	166	49	165	210
27.....	109	97	45	77	122	100	125	82	143	69
28.....	115	146	22	76	100	85	133	46	177	192
29.....	91	120	59	67	120	106	136	57	144	88
30.....	100	102	64	77	128	91	134	79	126	50
31.....	96	129	49	80	120	94	134	55	142	91
32.....	93	122	52	79	117	101	128	55	153	92
33.....	95	95	66	77	119	107	143	66	131	59
34.....	117	113	31	47	129	82	137	78	164	151
35.....	89	114	79	81	112	102	121	77	125	29
36.....	120	107	83	76	94	109	109	85	118	21
37.....	90	135	29	80	106	93	174	48	145	171
38.....	90	106	50	54	146	110	140	50	154	140
39.....	81	116	42	78	129	113	140	42	160	140
40.....	72	144	50	60	124	92	112	57	189	174
41.....	96	101	58	79	121	106	134	68	137	62
Total.....	4240	4757	1904	3147	4919	3932	5259	2803	5942	
Reduced sum of squares (hundreds).....	105	92	85	36	103	51	74	103	112	
Column minimum.....	70	89	22	47	88	66	105	42	115	

9 in trials 23 and 36; strain 5 exceeds strain 9 in trials 7, 12, and 30. Strain 7 exceeds strain 9 in trials 14, 26, 30, 33, and 37. Let us therefore consider the reduced matrix with strains 3 and 9 omitted.

Strain 8 is dominated by all strains except strain 3, which it dominates except for some scattered trials when strain 8 is not completely dominated.

However, the showing of strain 8 is so poor that we can assign it rank 8 without much error. Strain 4 appears to belong in rank 7, and strain 6 by the same considerations can be given rank 6. It is more difficult to assign ranks to 1, 2, 5, 7 which have many displacements in relative yielding ability. How-

TABLE 10
GAME MATRIX FOR GROWER VERSUS NATURE. NINE STRAINS OF
HANNCHEEN BARLEY, 1957-1958-1959. RESULTS IN PERCENT
OF TRIAL MEANS

Trial code	Kernel weight Grower (strains)									Reduced sum of squares (hundreds)
	1	2	3	4	5	6	7	8	9	
RB.....	96	120	85	92	106	106	101	85	110	11
RL.....	99	118	81	90	112	104	103	83	109	13
R2H.....	100	122	85	89	109	102	101	82	109	13
Rod row.....	101	123	92	92	107	101	100	85	99	9
PDB.....	100	118	90	88	109	101	102	81	111	11
PD2H.....	98	119	88	86	109	102	103	85	110	11
PDL.....	100	117	84	85	110	103	106	81	113	14
WL.....	96	117	84	93	111	100	104	88	107	9
W2H.....	86	126	84	97	114	104	97	83	109	17
WB.....	98	120	84	89	114	101	101	85	108	12
CB8(BAC).....	98	126	97	92	107	91	98	86	107	12
1959, XHR.....	97	121	102	83	107	100	100	83	106	11
S.....	106	114	99	85	106	104	102	78	106	10
Total.....	1275	1561	1155	1161	1421	1319	1315	1085	1404	
Reduced sum of squares (tens)	24	14	54	19	10	14	10	9	14	
Column minimum.....	86	114	84	83	106	91	97	78	99	

ever, consideration of the number of times each is dominant over the others gives strain 1 rank 5, strain 2 rank 4, strain 5 rank 3, and strain 7 rank 2. Thus, we arrive at the ranking 9, 7, 5, 2, 1, 6, 4, 8, 3; but strains 5 and 2 are so close as to be almost indistinguishable.

From the standpoint of odds estimated from the number of times each is dominant, we might grow strains 7, 5, 2, 1 in the ratio of 55:44:36:32, with the expectation of doing as well as possible against Nature with just these four strains. In general, we need to know the complete relative frequency distribution of Nature's strategies if we wish to compute accurate odds for the growing of the different strains. If we could predict Nature's strategy in advance then our choice might be narrowed to just one strain or a small number of strains, the particular choice being of no importance.

The only estimate we have of the long-run relative yielding ability of the strains are the column totals in table 9. It may be that these totals should be

weighted in some way, since the trials were deliberately designed to cover a very wide range of conditions, but it is not clear just how. On the basis of these totals, the ranking should be 9, 7, 5, 2, 1, 6, 4, 8, 3 with strain 9 having an average relative yield of about $2\frac{1}{2}$ times strain 3 and so on.

The whole structure of field trials is made clear by a contemplation of table 9. Sometimes one strain does better and sometimes another. Sometimes all strains are close together and sometimes they are far apart. Some strains are very erratic in relative yield and some are very consistent. Any recommendations for the use of certain strains and the proportions of their use must take account of these facts.

Game Matrix for Kernel Weights

The general aspects of table 10 are similar to those of table 9. If we rank the strains according to column minima, we get 2, 5, 9, 7, 6, 1, 4, 3, 8. The ranking by column totals is 2, 5, 9, 6, 7, 1, 4, 3, 8. Dominance considerations give a ranking of 2, 5, 9, 6, 7, 1, 4, 3, 8. It is obvious that choosing among sets of strains with different odds may give better results than choosing to grow one strain of the set exclusively.

For yield and kernel weights, we have the rankings based on column totals.

Rank	Yield	Kernel Weight
1	9	2
2	7	5
3	5	9
4	2	6
5	1	7
6	6	1
7	4	4
8	8	3
9	3	8

Now as a basis of recommendation, we must consider both of these ratings as well as the trial-to-trial variation in relative standing and possibly other factors. For instance, strain 2 is outstanding in kernel weight and rather consistent. Also, in yield it is quite high but variable. Because of kernel weight, strain 2 might be recommended over 7, 9, and 5. Also, strain 5 might be recommended over strains 9 and 7.

DISCUSSION

Snedecor (1956) and others have described mathematical models for field trials which depend on random assignment of varieties and replications in formal patterns. Johnson, Truoy, Rigney and Sprague (1954) feel that these models are so effective that the reporting of field trials can be reduced to a robot-like filling out of a standard form.

However, there is abundant evidence to show that the soil is a very complex manifestation of nature that seldom conforms to simple and rigid mathematical models that are easily analyzed by routine techniques. Areas of uniform response, identifiable only after growing, appear as islands which are irregularly distributed, not necessarily contiguous and usually include parts of several blocks, rows or columns.

Conventional designs, therefore, may give misleading results, because the "experimental error" is actually an indeterminate portion of soil variability which is unpredictable, even on apparently uniform soils, and is often extreme. It should also be noted that stress conditions vary from trial to trial and that under some stress conditions all varieties tested will do equally well while their response to other stress conditions may vary widely.

The reliability of conventional procedures, utilizing rigid mathematical models, has been questioned and discussed in detail by Baker and others (see "Literature Cited").

Hoyle (1959) developed a realistic background of environmental variation to show how stress indicators which are observable morphological plant characteristics can be used for grouping "varieties." Later, Hoyle and Baker (1959*a, b, c*; 1960) and Hoyle (1959) discussed in detail the essential island-like aspects of the variation of plant response to restricted environments and the use of this island concept in analysis of field trials.

The extent of the variation of apparently uniform soils with respect to essential chemical constituents and the implications for the accuracy of field trials were investigated and reported by Waynick (1919) and Waynick and Sharp (1919). Veihmeyer and Hendrickson (1938) have shown that soil moisture is also very nonuniform. Proebsting (1929) has indicated that the soil solution at a particular location varies markedly with time. The effect of all these variables on crop yield can be partially indicated by uniformity trials such as the one reported by Wiebe (1935) for grain yield of rod-row nursery plots for wheat.

The basic question investigated in the trials reported here is whether or not field trials can be successful independent of wide differences in environment and whether an unrealistic uniformity, if it could be achieved, would be necessary or desirable.

CONCLUSION

The ideal conditions upon which the usefulness of the classical field trials depend rarely exist in nature. Nor can the conditions that actually do exist be sufficiently formalized by current conventional practices to provide an adequate and consistently reliable interpretation of field trials.

Productivity levels show extreme variations within a relatively small environment and over short periods of time. Year-to-year conditions differ. Areas of uniform response cannot be identified in advance of planting, as recommended by classical procedures, but appear as irregular islands after growth. No so-called block or row effect has been found to exist in nature.

The experiments reported here involved no effort to find uniform environments in advance of planting but utilized the differing local conditions as they existed (including differences in soil and year grown) and intensified these differences by such devices as differences in seeding rate, plot size, and rate of irrigation.

The relative merits of the varieties tested were found to be stable over the full range of differences provided by time, place and procedure.

Raw data supplied by the tests were subjected to three different statistical treatments: factor analysis, game analysis and the SD analysis (developed

by the authors and based on the idea of stability of relative rank in like islands of variation). There was close correspondence of the rankings of the varieties provided by each of the three treatments.

Classical procedures employed for comparisons during the tests provided inaccurate or ambiguous results.

Field trials can be free of the procedures which attempt, by pre-selection and mathematical compensations, to arrange naturally irregular conditions into a convenient series of blocks or rows.

Simpler and more accurate procedures, based on conditions as they actually exist, can be substituted.

Thus the costs of definitive field trials can be greatly reduced. Time and labor can be saved. Agricultural testing programs can be accelerated, with increased confidence in the accuracy of the results.

LITERATURE CITED

BAKER, G. A.

1941. Fundamental distributions of errors for agricultural field trials. *National Mathematics Magazine* 16:7-19.

1944. F values for samples of four and four from populations which are the sum of two normal populations. *National Mathematics Magazine* 19:62-63.

1945. Test of the significance of the differences of per cents of emergence of seedlings in multiple field trials. *Journal of the American Statistical Association* 40:93-97.

1952a. Uniformity field trials when differences in fertility levels of subplots are not included in experimental error. *The Annals of Mathematical Statistics* 23:289-293.

1952b. Field trial problems. *The Annals of Mathematical Statistics* 23:480, Abstract.

BAKER, G. A., and R. E. BAKER

1953. Strawberry uniformity yield trials. *Biometrics* 9:412-421.

BAKER, G. A., and F. N. BRIGGS

1945. Wheat bunt field trials. *Journal of the American Society of Agronomy* 37:127-133.

1949. Wheat bunt field trials. II. Proceedings of the Berkeley Symposium on Mathematics and Probability 1945-46, pp. 485-491.

1950. Yield trials with back-cross derived lines of wheat. *Annals of the Institute of Statistical Mathematics* 2:61-67.

BAKER, G. A., and G. C. HANNA

1949. Transformations of split-plot yield trial data to improve analysis of variance. *Proceedings of the American Society for Horticultural Sciences* 53:273-275.

BAKER, G. A., M. R. HUBERTY, and F. J. VEIHMAYER

1952. A uniformity trial on unirrigated barley of ten years' duration. *Agronomy Journal* 44:267-270.

BAKER, G. A., and E. B. ROESSLER

1957. Implications of uniformity trial with small plots of wheat. *Hilgardia* 27:183-188.

BAKER, R. E., and G. A. BAKER

1950. Experimental design for studying resistance of strawberry varieties to *verticillium wilt*. *Phytopathology* 40:477-482.

CATTELL, R. B.

1952. *Factor Analysis*. Harper and Brothers, New York, xiii + 462 pages.

COLWELL, R. N.

1956. Determining the prevalence of certain cereal crop diseases by means of aerial photography. *Hilgardia* 26:223-286.

HANNA, G. C., and G. A. BAKER

1951. Analysis of asparagus field trials on the basis of partial records. *Proceedings of the American Society for Horticultural Sciences* 57:273-276.

HOYLE, B. J.

1957. Stress indicators. Annual Research Review of the Agronomy Staff, Davis, January 28.

1958. Islands. Annual Research Review of the Agronomy Staff, Davis, January 28.

1959. Tulelake Field Station 13th Annual Report—Research. Mimeograph, 20 pages.

HOYLE, B. J., and G. A. BAKER

1959a. A more efficient barley yield testing method. Barley News Letter 2:32-33.

1959b. The S.D. technique and its use for conducting field trials. Four pages. Mimeographed.

1959c. The analysis of field trials based on the concept of islands of variation. Agronomy Abstracts, American Society of Agronomy, Cleveland Meetings, Nov. 16-20.

1960. Factor analysis of twenty-eight independent field trials on nine strains of Hannechen barley. Biometrics 16:127-128. Abstract No. 633.

HOLZINGER, K. J., and HARRY H. HARMAN

1941. Factor Analysis. University of Chicago Press, Chicago, xii + 417 pages.

KENDALL, M. G.

1957. A course in Multivariate Analysis. Hafner Publishing Company, New York, 185 pages.

JOHNSON, I. J., EMIL TRUOY, J. A. RIGNEY, and GEORGE SPRAGUE

1954. Publication of analysis of variance tables. Agronomy Journal 46:475.

PROEBSTING, E. L.

1929. Changes in the nitrate and sulfate content of the soil solution under orchard conditions. Hilgardia 4:57-76.

RIDDLE, O. C., and G. A. BAKER

1944. Biases encountered in large-scale yield tests. Hilgardia 16:1-14.

SMITH, F. L.

1858. Effects of plot size, plot shape, and number of replications on the efficiency of bean yield trials. Hilgardia 28:43-63.

SNEDECOR, GEORGE W.

1956. Statistical Methods. Fifth Edition. Iowa State College Press, Ames, xiii + 534 pages, especially 251-253.

THURSTONE, L. L.

1947. Multiple-Factor Analysis. University of Chicago Press, Chicago, xix + 535 pages.

VEIHMEYER, F. J., and A. H. HENDRICKSON

1938. Soil moisture as an indication of root distribution in deciduous orchards. Plant Physiology 13:169-177.

WAYNICK, D. D.

1919. Variability in soils and its significance to past and future soil investigations. I. A statistical study of nitrification in soils. University of California Publications in Agricultural Sciences 3:243-270.

WAYNICK, D. D., and L. T. SHARP

1919. Variability in soils and its significance to past and future soil investigations. II. Variations in nitrogen and carbon in field soils and their relation to the accuracy of field trials. University of California Publications in Agricultural Sciences 4:121-139.

WIEBE, G. A.

1935. Variation and correlation in grain yield among 1,500 wheat nursery plots. Journal of Agricultural Research 50:331-357.

WILLIAMS, J. D.

1954. The Compleat Strategyst. Rand Series. McGraw-Hill, New York, xiii + 234 pages.

APPENDIX A

Computation of the SD Score

In the comparison of varieties by a field trial we must first select a plot of land (*P*) and assign the varieties to be compared to certain sub-plots. A very effective arrangement is the BH design shown in appendix fig. 1.

APPENDIX FIG. 1

The BH Design

7	6	2	1	3	5	6	8	4	6
1	2	8	3	6	4	9	5	7	3
7	3	4	5	2	9	1	6	8	9
5	6	9	7	1	8	4	2	3	6
2	7	6	4	9	1	3	8	5	1
3	4	5	8	7	2	6	9	1	5
9	8	1	6	5	3	2	7	4	1
6	5	2	1	4	7	8	3	9	5
8	1	7	9	3	6	5	4	2	9
4	9	3	2	8	5	7	1	6	7

Appendix Fig. 1. The area inside the double line is a basic 9 by 9 latin square and all varieties are paired at least once. By including the extreme top row and right column all varieties are paired twice or more. In addition to these features all varieties are included once in each of the 9 small blocks enclosed by the single lines. This design provides extreme uniformity of coverage for the plot area. It may be modified for more or fewer varieties.

Table of descending order of observed yields.

Having grown and harvested the plot, the sample values for the desired attribute are listed in descending order for each variety. Since each sample value of a variety's range represents a different level of productivity within *P*, then the highest values of each range have obviously come from a sub-area within *P* of the highest, relative, productivity level. The lowest values of each variety range have come from the sub-area of least productivity for *P*, and the intermediate values from productivity levels of intermediate value. The very best variety comparisons possible, then, are those made by comparing all varieties at *each* level of productivity. This is not the same as comparing the mean variety values as will be brought out.

Observe Appendix Table 1 for an illustration of the SD analysis. In this example, single, one-foot rows were used. For ease of reference, each level of productivity will be called an "island" and the number of islands in any *P* is assumed to be equal to the number of samples taken per variety. These islands are represented by the 9 horizontal rows of Appendix Table 1.

SD Table.

Each row in the table of descending values contains one value from each variety. Row 1 includes the highest value for each variety, row 2 the second highest value for each variety, etc. Starting with the first row, the varieties are ranked by name and entered as names in the first column of the SD Table

(Appendix Table 2). The second row is ranked similarly and entered as variety names in the second column, etc. Each row in the table of descending values and each corresponding column in the SD Table constitute the sample values inclusive within one island of variation. There will be as many columns in the SD Table as there are replicated values. Assuming 9 varieties and 9 replications the SD Table would have 9 columns; each of which show the ranking order by name for each row of the table of descending values. Three more columns are added to complete this table. Column 10 is a ranking of the mean values of each variety as determined from Appendix Table 1, while column 11 is made up of the variety names corresponding to the values of column 10. Column 12 is reserved for entering the SD score when computed.

Computing the SD Score.

The SD score is a number which signifies the stability, or agreement between the ranking order of a given variety in the column of ranked mean values with each column of island values. In our example there are 9 chances for a variety to agree or disagree in its rank to that of its mean. Each instance where there is agreement a value of 0 is assigned. Each instance where the agreement is missed by one rank in either direction a value of 1 is assigned, and in each instance where the agreement is missed by 2 or more, a value equal to the base number is assigned. The base number is equal to (the number of sample values) + 1, or, in this case, 10.

An Example.

In Appendix Table 2 there are 9 varieties and 9 replicate values. Variety 8 ranks 4th highest according to its mean value. Among the 9 islands there are 2 cases where variety 8 deviates by one rank from the fourth position, one case where it deviates by more than one rank from the fourth position, and in the remainder of the islands it ranks fourth. The SD score is $1 + 1 + 10 = 12$.

Reconstruction.

Suppose we are given the mean yields of 9 varieties, their SD scores, and the base number (10 in this case). For the score of 12, divide by the base number and obtain the whole number 1 with a remainder of 2. The whole number is always equal to the deviations of over one rank, and the remainder is always equal to the number of deviations of one rank only. The number of islands in agreement is found by subtraction.

For a variety with a score of 7 and base number 10, divide 7 by 10 and obtain the whole number 0 and remainder 7 which indicates that there were no deviations of over one rank, and 7 out of 9 islands had deviations of one rank only.

Applying the SD Score to Groups.

In many examples the SD score is applied to a group of varieties instead of a single one. With 9 varieties it may be that the varieties can be grouped as: (1 and 2), (3, 4, and 5), and (6, 7, 8, and 9). The scoring is identical except that a group is considered instead of an individual. Groups are determined by inspection and members within a group are usually considered equal to each other, but different from the other groups.

General Remarks.

The SD score is applicable to the individual variety mean and tells at a glance the stability it has. Only those means with a score below a "significant" level are considered precise enough for comparing with other plot means. One deviation of *over one rank* and four deviations by *one rank only* (score 14) per 9 samples has been found to agree very well with observations. For those sample means with 'significant' scores one may be said to be better than the other by the numerical difference indicated for the plots concerned.

Many pattern types occur in SD tables. These each have specific meanings which the experimenter soon comes to recognize.

The interpretation obtained by the SD procedure outlined here has been rigorously tested and compared exhaustively with that obtained by the analysis of variance. The SD interpretation has been found to agree with the observed facts in all cases much more closely than the analysis of variance. Since the goal of any test plot is to determine which variety is best, such an answer is superior to a probability figure, as such.

One comparison quickly made of the effectiveness of the SD score is shown in Appendix Table 3. For the actual plot trials we would expect low scores, indicating variety differences, and from the uniformity trials we would expect high scores. This expectation holds generally true in each case.

The high SD scores in Appendix Table 4 indicates a random ranking of the treatments within the various islands of productivity. This is to be expected from the theory. Occasionally a dummy treatment from a uniformity trial is indicated as significant, as in experiments 10 and 11. This must be expected *when* the dummy design actually makes selections which are significant. However, the repeating of the whole experiment at least twice quickly dispels the possibility of a misinterpretation in a bona-fide field trial.

APPENDIX TABLE 1
VARIETY YIELDS IN GRAMS PER ONE-FOOT ROWS ARRANGED IN
DESCENDING ORDER

Island	Variety								
	1	2	3	4	5	6	7	8	9
1.....	62	91	38	69	109	71	90	49	106
2.....	61	89	33	55	99	70	88	42	105
3.....	60	86	30	49	97	67	87	41	102
4.....	59	85	29	47	95	65	86	41	102
5.....	57	84	22	45	91	58	83	41	96
6.....	54	83	20	44	89	53	82	33	89
7.....	49	74	18	44	81	51	81	32	82
8.....	48	71	17	32	80	45	66	22	78
9.....	38	68	12	28	79	44	63	20	75
Average.....	54.6	81.3	24.5	45.5	91.2	58.4	80.6	35.6	92.7

For island no. 1, variety 5 ranks highest, variety 9 next highest, and the other varieties follow in the order of 2, 7, 6, 4, 1, 8 and 3. This order is shown in the first column of table 2.

APPENDIX TABLE 2
SD TABLE
VARIETIES ARRANGED IN DESCENDING ORDER ACCORDING TO
ISLAND NUMBER

Rank	Island number									Variety		SD score
	1	2	3	4	5	6	7	8	9	No.	Aver.	
1.....	5	9	9	9	9	9	9	5	5	9	92.7	3
2.....	9	5	5	5	5	5	7	9	9	5	91.2	4
3.....	2	2	7	7	2	2	5	2	2	2	81.3	3
4.....	7	7	2	2	7	7	2	7	7	7	80.6	12
5.....	6	6	6	6	6	1	6	1	6	6	58.4	2
6.....	4	1	1	1	1	6	1	6	1	1	54.6	3
7.....	1	4	4	4	4	4	4	4	4	4	45.5	1
8.....	8	8	8	8	8	8	8	8	8	8	35.6	0
9.....	3	3	3	3	3	3	3	3	3	3	24.5	0

L.S.D. for 1 percent level is 12.5 as calculated for the actual field arrangement.

SD scoring method:

Variety agreement with the mean rank scores 0

Each deviation by one rank scores 1

Each deviation by over one rank

scores $10 = (n + 1)$.

A "significant" score is considered as 14 for this example, that is, $(n + 1) + [(1/2)n]$ rounded down.

Significance merely means that for this plot the ranking was stable enough to consider it for determining a real difference.

APPENDIX TABLE 3
SD SCORES OF ACTUAL PLOT VARIETY TRIALS UTILIZING THE
BH DESIGN
(varieties not designated but are ranked in descending order)

Rank	Experiment number											
	1	2	3	4	5	6	7	8	9	10	11	12
1.....	0	2	10	4	0	3	2	12	0	1	4	3
2.....	0	5	4	4	18	4	14	9	22	36	15	4
3.....	0	4	4	5	1	1	4	16	13	26	3	3
4.....	1	3	10	5	11	6	39	16	28	1	5	11
5.....	1	19	0	0	12	6	20	7	3	27	3	2
6.....	0	0	1	13	20	2	2	2	11	6	5	3
7.....	0	0	2	5	5	2	20	1	0	12	24	1
8.....	0	0	1	1	5	2	20	0	0	12	2	0
9.....	0	0	0	0	9	2	0	0	0	1	0	0

Significant level is 14 or lower.
Each column contains the SD scores for the same 9 barley varieties. Each experiment was of a different plot size or cultural practice. Varieties do not necessarily rank the same for each experiment, but do so very closely. The high number of significant SD scores indicates real variety differences.

APPENDIX TABLE 4
SD SCORES FROM UNIFORMITY TRIALS WITH DUMMY TREATMENTS
SELECTED BY USE OF BH DESIGN

1	32	30	40	60	41	26	45	29	32	61	32	44
2	35	44	74	43	26	35	23	23	65	26	44	33
3	25	34	65	53	44	42	48	32	73	36	34	71
4	36	53	81	62	63	35	65	48	23	52	33	42
5	45	43	37	72	81	53	20	30	41	35	32	32
6	43	25	40	80	36	53	47	48	63	63	24	32
7	44	27	48	35	33	54	38	40	65	44	5	42
8	33	35	39	62	22	45	19	14	29	7	21	34
9	50	60	45	60	12	62	36	19	65	6	21	11

Significant level is 14 or lower.
The first 9 experiments are from bean trials (Smith, F. L., 1958), and the last three from strawberry plots Baker, G. A., and Baker, R. E., 1953).

